QUARTERLY PROGRESS REPORT

No. 83

OCTOBER 15, 1966

NASA

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS
CAMBRIDGE, MASSACHUSETTS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS

QUARTERLY PROGRESS REPORT No. 83

October 15, 1966

Submitted by: H. J. Zimmermann
G. G. Harvey

TABLE OF CONTENTS

CONTENTS

CONTENTS

PERSONNEL

Administration

Prof. H. J. Zimmermann, Director
Prof. G. G. Harvey, Associate Director
Mr. R. A. Sayers, Assitant Director

Advisory Committee

Dean G. S. Brown
Pfor. W. W. Buechner
Prof. W. B. Davenport, Jr.
Prof. P. Elias
Prof. G. G. Harvey
Prof. A. G. Hill
Prof. I. W. Sizer
Prof. H. J. Zimmermann
(Chairman)

Research Committee

Dean S. C. Brown
Prof. L. J. Chu
Prof. M. Halle
Prof. G. G. Harvey
Prof. W. A. Rosenblith
Mr. R. A. Sayers
Prof. W. M. Siebert
Prof. L. D. Smullin
Prof. M. W. P. Strandberg
Prof. P. D. Wall
Prof. J. R. Zacharias
Prof. H. J. Zimmermann
(Chairman)

Professors

Allis, W. P.	Haus, H. A.	Schreiber, W. F.
Barrett, A. H.	Hill, A. G.	Shannon, C. E.
Bitter, F.	Huffman, D. A.	Shu, G. G. (Visiting)
Bose, A. G.	Ingard, K. U.	Siebert, W. M.
Brown, S. C.	Jakobson, R.	Sledd, J. (Visiting)
Burke, B. F.	Kerrebrock, J. L.	Smullin, L. D.
Chomsky, N. A.	King, J. G.	Stevens, K. N.
Chu, L. J.	Kurylowicz, J. (Visiting)	Strandberg, M. W. P.
Eden, M.	Lee, Y. W.	Wall, P. D.
Edgerton, H. E.	Lettvin, J. Y.	Warren, B. E.
Elias, P.	Mason, S. J.	Waugh, J. S.
Gyftopoulos, E. P.	Minsky, M. L.	Wozencraft, J. M. (Absent)
Halle, M.	Rose, D. J.	Zacharias, J. R.
Harvey, G. G.	Rosenblith, W. A.	Zimmermann, H. J.

Associate Professors

Bekefi, G.	Hennie, F. C.	Loewenthal, M.
Bers, A.	Hoffman, M. A.	Massey, J. L. (Visiting)
Brown, G. A.	Jackson, W. D. (Absent)	Matthews, G. H.
Dennis, J. B.	Katz, J. J.	McCune, J. E.
Dupree, T. H.	Kleppner, D.	Mortenson, E. (Visiting)
Fodor, J. A.	Klima, E. S.	Oates, G. C.
Gallager, R. G.	Kyhl, R. L.	Peake, W. T.
Garland, C. W.	Lee, F. F.	Pomorska, Krystyna

PERSONNEL

Associate Professors (continued)

Rafuse, R. P.
Searle, C. L.

Taylor, E. F. (Visiting)
Teager, H. M.

Van Trees, H. L.
Watkins, C. (Visiting)

Assistant Professors

Anderson, J. (1)
Bernard, G. D. (1)
Billman, K. W.
Black, W. L. (1)
Blum, M.
Bowers, K. W.
Briggs, R. J.
Brown, J. E.
Bruce, J. D.
Carabateas, E. N. (Absent)
Cheng, H.
Dean, L. W., III
Fiocco, G.

Goutmann, M. M.
Hoversten, E. V. (1)
Huang, T. S.
Ingraham, J. C.
Katona, P. G. (1)
Kennedy, R. S.
Kinsey, J. L.
Kiparsky, R. P. V.
Klatt, D. H. (1)
Lee, H. B.
Lenoir, W. B. (1)
Lidsky, L. M.
Nelsen, D. E.

Oppenheim, A. V. (1)
Perry, C. H.
Pomeranz, B.
Siambis, J. G. (1)
Snyder, D. L.
Spann, R. N.
Staelin, D. H. (1)
Stickney, R. E.
Troxel, D. E.
Wagner, T. J. (Visiting)
Weiss, R.
Weiss, T. F.
Yip, S.

Lecturers

Ferretti, E.
Mark, R. G.

Nedzelnitsky, V.
Pitts, W. H.

Instructors

Blesser, B.
Burns, S. K.
Eisenberg, M.
Evans, J. E.
Henke, W. L.

Kukolich, S. G.
Kuroda, S-Y.
McEnally, T. E., Jr.
Parker, R. R.
Prabhu, V. K.

Schaefer, R. W.
Schindall, J. E.
Schneider, H. M.
Speck, C. E.
Tripp, A. P.

Research Associates

Barnett, G. O.
Bromberger, S.
Canfield, J. V.
Chung, K.
Durlach, N. I.
Garrett, M. F. (2)

Grams, G. W.
Hall, R. D.
Hoffman, R. A.
Kittelburger, J. S.
Kolers, P. A.
Kornacker, K.

Novotny, D. B.
Papert, S. A.
Sears, R. E. J.
Smith, T. G. Jr.
Steinbrecher, D. H.
Zisk, S. H.

Guests

Bailey, C. J. M.
Bullowa, Margaret
Burnham, D. C.
Chang, Teresa
Fintoft, K.
da Fonseca, J. L. S.

Fraser, J. B.
Gaines, B. E.
Kessler, A. R.
Myint, M. T.
Nevo, E.

Palma-Vittorelli, Maria B.
Pritchard, D. E.
Sharda Nand
Walker, D. E.
Williamson, K. R. M.
Winkler, P. F.

(1) Engineering Postdoctoral Fellow   (2) NIH Postdoctoral Trainee

PERSONNEL

Visiting Scientists

Moreno-Diaz, R.
Nomoto, M. (1)
Suzuki, J.

Research Affiliates

Barlow, J. S.
Brodey, W. M.
Brown, R. M.

Crist, A. H.
Fohl, T.
Howland, B. (2)
Langbein, D.

Mayo, J. W.
McLardy, T.
Ozier, I.

Postdoctoral Fellows

Borbely, A. A.
Charney, Elinor K.
Fokkens, N.
Franzen, O.
Gruber, J. S. (3)

Hartman, H. (4)
Hellekant, B. C.
Lampis, G. (5)
Milner, J-C. G.

Natapoff, A. (4)
Schwartz, A. (4)
Smith, N. V.
Songster, G. F. (4)
Taub, A. (4)

R. L. E. Research Staff

Barrett, J. W.
Benhaim, N.
Burgess, R. G.
Cattell, N. R.
Chung, S-H.
Crowther, Patricia P.
Cunningham, A. W. B.
Davis, Heather S.
Edwards, D. J.
Fischler, H.
Fontaine, C. L.
Frater, Gail M.
Gambardella, G.
Harwitt, Joan
Ingersoll, J. G.

Ingham, K. R.
Isaacs, E. C.
Jensen, E. R.
Johnston, W. D., Jr.
Kiang, N. Y. S.
Kierstead, J. D.
Kim, C-W.
Mattison, E. M.
McCaffrey, A.
McCarthy, J. J.
McCulloch, W. S.
McLoud, Veronica E.
Menyuk, Paula
Mulligan, W. J.
Novenski, A. F.
O'Brien, F. J.

Palma, M. U.
Pennell, Martha M.
Perkell, J. S.
Pitts, W. H.
Plummer, W. W.
Portinari, J. C.
River, Eleanor C.
Rojas-Carona, R. R.
Rosebury, F.
Ryan, L. W.
Shaw, M. L.
Tretiak, O. J.
Viertel, J. J.
Wawzonek, J. J.
Wickelgren, G. L.

Research Assistants

Austin, M. E.
Baggeroer, A. B.
Bartsch, R. R.
Brown, T. S.
Catto, P. J.
Chan, S. W-C.
Chandra, A. N.
Chase, D.
Chen, K. R-S.
Ching, H.

Chou, S.
Citron, A.
Davis, J. A.
Dean, Janet P.
DeRijk, R. P. G.
DeWolf, J. B.
English, R. P.
Ezekiel, S.
Fehrs, D. L.
Fisher, C. H.

Flynn, R. W.
Frediani, J. K.
Gabrielian, A.
Gadzuk, J. W.
Gustafson, K. T.
Ham, D. O.
Herba, F.
Hillman, A. F., Jr.
Hoff, P. W.
Hougen, M. L.

Research Assistants (continued)

Huang, T.
Kirk, R.
Kitrosser, D. F.
Klouman, P. H. B.
Klumpp, M. H.
Koons, H. C.
Kurth, R. R.
Kusse, B. R.
Levy, E. K.
Liu, J-H.
Logan, R. M.
Lou, D. Y-S.
Makhoul, J. J.
Mangano, J. A.
Maul, M. K.

Melnick, M.
Milne, D. C.
Moir, R. W.
Moldon, J. C.
Moran, J. M. Jr.
Moses, J.
Murakami, M.
Nahvi, M.
Ng, L. C.
Oates, D. E.
Offenberger, A. A.
Papadopoulos, G. D.
Perozek, D. M.
Poulo, L. R.
Poussart, D. J-M.

Richters, J. S.
Rogers, A. E. E.
Ross, A. H. M.
Snyder, D. D.
Sugawara, A.
Tomlinson, R. S.
Wagner, C. E.
Wallace, R. N.
Weidner, M. Y.
Weiner, B. B.
Williams, J. A.
Wong, Y-M.
Yamamoto, S.
Young, R. A.
Zeiders, G. W. , Jr.

Graduate Assistants

Andrews, M. L.
Brown, T. R.
Davis, W. B.
Ewing, H.
Fertel, Jeanne H.
Forrester, V. G.
Fox, R. L.
Free, J. U., Jr.

Garosi, G. A.
Golub, R.
Gschwendtner, A. B.
Guttreich, G. L.
Langdon, R. M., Jr.
Llewellyn-Jones, D. T.
McClintock, J. E.
Pleasance, L. D.

Reifenstein, E. C., III
Reznek, S. R.
Silk, J. K.
Smith, T. B.
Vellenga, J. H.
Wilheit, T. T., Jr.
Wright, B. L.
Yung, B. N.

Teaching Assistants

Anderson, G. B.
Bice, P. K.
Columbant, D. G.
Crane, D. E.
Dum, C. T.
Freeman, J. A. (1)
Glaser, J.
Greenwood, R. E.
Guttman, D. S.
Harris, R. V., III

Hartmann, H. P.
Hill, R. A.
Khanna, M.
Kinsley, R. W., Jr.
Kososki, J. F.
Lazarus, M. B.
Levin, M. I.
Metz, P. J.
Meyn, J. H.

Moxon, E. C.
Mozzi, C.
Peters, P. S., Jr. (2)
Portner, E. M., Jr.
Samis, M. A.
Schaefer, D. W.
Seitz, C. L.
Singer, J. J.
Stone, E. T.
Wawzonek, J. J.

Graduate Students

Akmajian, A. (3)
Allen, J. L. (6)
Allen, R. J. (4)
Anderson, J. A. (2)
Anderson, S. R. (5)
Arnstein, D. S. (6)
Barnwell, T. P., III (8)
Balcewitz, J. F. (7)

Bedell, G. D., IV (3)
Bever, T. G. (9)
Bhushnan, A. K. (10)
Bigham, T. D., Jr. (6)
von Bismarck, G.
Blum, G. D. (8)
Bowers, J. S. (3)
Braida, L. D. (8)

Brame, M. K. (10)
Browne, E. W. (3)
Brueck, S. R. J. (8)
Bucher, E. A. (8)
Cain C. A. (10)
Caldwell, D. (8)
Callen, J. D. (11)
Carter, R. J. (2)

(1) United States Air Force
(2) NIH Trainee
(3) National Defense Education Act Fellow
(4) RCA Fellow
(5) NSF Trainee
(6) Lincoln Laboratory Associate

(7) Jonathan Whitney Fellow
(8) NSF Fellow
(9) Harvard Fellow
(10) Bell Telephone Laboratories Fellow
(11) U.S. AEC Fellow

Graduate Students (continued)

Chapin, P. G. (2)	Kolodny, Nancy H. (3)	Rizk, H. M. (24)
Clark, N. A. (3)	Koskinen, M. F.	Roberson, J. E. (6)
Colburn, H. S. (4)	Krakauer, L. J. (17)	Rogoff, G. L.
Collins, L. D. (3)	Lackner, J. R. (5)	Rolland A. E. (23)
Cruise, T. J. (3)	Lawter, J. R. (3)	Ross, J. A. (3)
Culicover, P. W. (5)	Leonardi-Cattolica, A. M. (17)	Ross, J. R. (1)
Davis, A. M. (6)	Lopez, O. (10)	Saleh, A. M. (24)
Davis, J. A. (4)	Lubin, M. D.	Schulz, H. M., III (3)
Decher, R. (7)	Manheimer, W. M. (3)	Sheinson, R. S.
Dougherty, R. C. (1)	Max J. (13)	Shupe, D. S. (3)
Edmonds, J. E. (1)	McDowell, G. Q. (3)	Simpson, J. I. (3)
Fetz, E. E. (3)	McNary, C. A. (3)	Snow, M. S. (3)
Fidelholtz, J. L. (1)	Mendell, L. M. (2)	Spielman, C. A. (5)
Flannery, D. L. (3)	Merrill, E. G. (18)	Stanley, R. J. (1)
Gaut, N. E. (7)	Mildonian, A. A., Jr. (6)	Stephenson, R. S. (10)
Geis, M. L. (8)	Mozzi, R. L. (19)	Strong, R. M. (25)
George, E. V. (9)	Muehlner, D. J. (3)	Suchard, S. N.
Goldberg, A. J. (10)	Mueller, P. E. (20)	Swain, D. W. (26)
Goldfield, R. (5)	Myers, Amy E. (5)	Tate, R. H. (6)
Good, W. E. (11)	Naro, A. J. (1)	Thiersch, C. L. (3)
Greaves, W. (16)	Nelson, A. C.	Thomae, I. H. (2)
Greenspan, R. L.	Nelson, G. P. (12)	Thomas, R. H. (7)
Gross, L. N. (1)	Nolan, J. J., Jr.	Tornberg, N. E.
Guinan, J. J., Jr. (3)	Odette, G. (15)	Tremain, R. E., Jr.
Guldi, R. L. (10)	Parrish, J. H. (3)	Vugrinec, Z. (27)
Harris, J. W. (12)	Perlmutter, D. M. (5)	Walker, J. L. (23)
Hartline, D. K. (12)	Peterson, D. L.	Wallace, R. N. (3)
Heggerstad, H. M. (13)	Peterson, R. E.	Walpert, G. A. (10)
Heller, J. A. (14)	Pilc, R. (6)	Wang, C. H. (21)
Hofmann, T. R. (1)	Pinkston, J. T., III (3)	Weinstein, C. J. (10)
Howard, I. J. (5)	Pittenger, L. C. (15)	Wiederhold, M. L. (12)
Hsiao, H. S.	Rabiner, L. R. (3)	Wilson, T. L. (3)
Hudis, M. (15)	Rack, H. J. (20)	Wolaver, D. H. (3)
Jackendoff, R. S. (5)	Ramshaw, J. D. (3)	Wold, J. J. (3)
Jameson, P. W.	Raymond, S. A. (10)	Woo, Nancy H. (1)
Katyl, R. H. (3)	Reeves, J. P. (1)	Wright, D. A. (9)
Kayne, R. S. (5)	Ritter, J. T. (21)	Wright, W. A.
Kimball, J. P. (1)		Wu, W. W.

(1) NIH Trainee
(2) Danforth Foundation Fellow
(3) NSF Fellow
(4) Sperry-Rand Fellow
(5) National Defense Education Act Fellow
(6) Bell Telephone Laboratories Fellow
(7) NASA Fellow
(8) Hughes Aircraft Fellow
(9) RCA Fellow
(10) NSF Trainee
(11) U.S. Rubber Corporation Fellow
(12) NIH Fellow
(13) Lincoln Laboratory Associate
(14) Xerox Fellow

(15) U.S. AEC Fellow
(16) Kennedy Fellow
(17) NSF Co-op Fellow
(18) Public Health Service Trainee
(19) Raytheon Fellow
(20) Vanadium Company Fellow
(21) Woodrow Wilson Fellow
(22) International A.E.C. Fellow
(23) United States Coast Guard
(24) U.A.R. Scholarship
(25) Hertz Foundation Fellow
(26) NASA Trainee
(27) United Nations Fellow

PERSONNEL

Undergraduates (Thesis or Special Problems)

Babitch, D.
Chase, D. L.
Costa, R. A.
Davidow, J. E.
DeBonte, R. J.
Eckstein, P. F.
Franz, J. M.
Galiger, P. E.
Gill, T. A.
Goodmann, J. M.
Jones, R. L.

Juvkam-Wold, H. C.
King, P. A.
Kolb, C. E., Jr.
Leary, A. R.
Leighninger, D. M.
Malpeli, J. G.
Mauer, J. L., IV
McNichols, L. A.
Memishian, J., Jr.
Morgan, H. D.
Partridge, L. D.
Patterson, J.

Pitegoff, A. D.
Sevick, G. E.
Smith, R. R.
Solin, J. R.
Strand, T. F.
Swobada, W.
Wandzelak, M.
Warshaw, A. S.
Wheeler, G. M.
Williams, F. K.
Wolfe, P. D.

Student Employees

Ackerman, W. B.
Bosel, J. P.
Engle, R. H.
Granek, H.
Howell, R. P.
Hung, H. L.
Jastrem, R. J., Jr.
Johnson, L. G.
Kinnaman, W. A.

Koehler, R. F.
Law, Sara
Li, E.
Marandino, G. F.
Michel, A.
Mitchell, M. M.
Naqvi, A. A.
Plice, W. A.
Postol, T. A.

Ray, J. N.
Reintjes, J. F.
Rich, L. J.
Ruderman, G. S.
Solarz, R. W.
Spalding, J. W.
Toong, H-M. D.
Waters, J. W.
Westerfeld, E. C.

R. L. E. Administrative Staff

Duffy, D. F.
Hewitt, J. H.
Keyes, R. V., Jr.

Sayers, R. A.
Smith, P. L.
Thomas, Helen L.

Administrative Assistant

Bella, C. J.

Office Clerks

Barron, Gladys G.
Butler, D. F.
Chase, Arbella P.
Engler, R. R.

Gregor, C. A.
Ippolito, Dorothy A.
O'Toole, J. P.
Peck, J. S.

Ruggere, P. A.
Scalleri, Mary B.
Simon, E.
Toebes, Rita K.

Typists

Greene, Yvonne E.
Murphy, Mary R.

Myers, Alberta L.
Smith, Mary L.

Stagliola, Eleanor E.
Young, Nancy E.

PERSONNEL

Technical Typists

Jones, Elizabeth
Kliphan, Evelyn C.

Mullin, Priscilla A.
Schröder, Gertraude L.

Secretaries

Baumann, Irene H. E.
Bedrosian, Isabel
Blum, Elaine
Carbone, Angelina
Cataldo, Donna L.
Cohen, Phyllis J.
Conwicke, Vera
Cosgrove, Mary J.
Cummings, Jane F.
DiPietro, Toni R.
Fernandez, Teresita M.

Geller, Elaine J.
Hamilton, Martha C.
Holden, Meredith F.
Hurvitz, Rose S.
Imperato, Eleanor M.
Ingersoll, Nancy L. B.
Johnson, Barbara A.
Kaloyanides, Venetia
Kirk, Marshall A.
Loeb, Charlotte G.
McCarthy, Barbara L.
McEntee, Doris C.

Owens, Mary E.
Paige, Okella M.
Parrella, Cynthia A.
Reid, Gloria C.
Ricker, Barbara J.
Skewes, Anne
Smith, Clare F.
Tomlinson, Ann B.
Tortici, Camille A.
Van Wezel, Ruth
Wanner, Patricia A.

Engineering Assistants

Berg, A. E.
Crist, F. X.

McKenzie, J. A.

Papa, D. C.
Thompson, J. B.

Technical Assistants

Byers, F. H.
Grande, Esther D.
Hall, Kyra M.
Iverson, Alice I.
Major, Diane

Newman, Charlotte M.
Rabin, Sylvia G.
Rosenthal, Kathryn F.
Salacinski, Barbara A.

Sanders, D.
Schumaker, N. E.
Seymour, P. L.
Swenson, Judith E.
Yaffee, M. A.

Technicians

Andrews, G. T.
Aucella, Alice G.
Barrows, F. W.
Beaton, Catherine M.
Butler, R. E., Jr.
Connolly, J. T.
DiPietro, P. J.
Ferragamo, A. A.
Fitzgerald, E. W., Jr.
Gage, R. B., Jr.

Gay, H. D.
Hill, R. F.
Iovine, M. A.
Kaufman, D. E.
Kelly, M. A.
Lewis, R. R.
MacDonald, K. R.
McLean, J. J.
Neal, R. W.
North, D. K.

Reid, E.
Schwabe, W. J.
Sears, A. R.
Sholder, J. A.
Sprague, L. E.
Stevens, J. A.
Summers, C. L.
Thibodeau, D. S.
Tortolano, A. J.
Wentworth, A. G., Jr.

Technicians' Shop

Lorden, G. J., Foreman

Fownes, Marilyn R.

Laboratory Assistants

Barrows, Helene G.
Cardia, P. F.
Miller, S. A.

PERSONNEL

Drafting Room

Navedonsky, C. P., Foreman
Donahue, J. B.

Hillier, Anna M.

Porter, Jean M.
Rollins, I. E.

Photographic Shop

Aradi, M. G.
Cook, J. F.
Karas, P.

Machine Shop

Keefe, J. B., Foreman
Aalrud, R. W.
Bletzer, P. W.
Brennan, J.
Bunick, F. J.

Cabral, M., Jr.
Carter, C. E.
Harvey, A. O.
Liljeholm, F. H.
Muse, W. J.

Reiman, W.
Ridge, P. A.
Ryan, J. F.
Sanroma, J. B.
Shmid, E.

Tube Laboratory

Rosebury, F.

Leach, G. H., Jr.
MacDonald, A. A.

Ryan, L. W.

Glass Shop

DiGiacomo, R. M.
Doucette, W. F.

Stock Clerks

Doherty, R. H.

Haggerty, R. H.
Legier, D. O.

Sharib, G.

Utility and Maintenance

Doiron, E. J., Foreman
Audette, A. G.

Lucas, W. G.
McDermott, J. F.

Riley, J. F.
Sincuk, J., Jr.

PUBLICATIONS AND REPORTS
MEETING PAPERS PRESENTED

Optical Society of America, Fiftieth Anniversary Meeting, Washington, D.C.
March 15-18, 1966

R. S. Kennedy and E. V. Hoversten, A Model of Propagation through a Turbulent Atmosphere for Optical Communication

American Physical Society Meeting, Washington, D.C.
April 25-28, 1966

C. O. Thornburg, Jr. and J. G. King, Search for an Electric Dipole Moment in the Cesium Atom

Seventy-first Meeting, Acoustical Society of America, Boston, Massachusetts
June 1-4, 1966

T. H. Crystal, Further Experiments Using a Larynx Model

N. Y. S. Kiang, J. W. Shipley, and M. B. Sachs, Sensitivity of Single Auditory Nerve Fibers to Acoustic Stimuli

W. T. Peake and J. J. Guinan, Jr., A Circuit Model for Transfer Characteristics of the Cat's Middle Ear

R. R. Pfeiffer, Input-Output Relations of Neurons in the Anteroventral Cochlear Nucleus Based on Component Analysis of the Waveforms

M. B. Sachs and N. Y. S. Kiang, Iso-rate Contours for Single Auditory Nerve Fibers

Lecture Series, Acoustics Institute, Technische Universität Berlin, Berlin, West Germany
June 12-July 19, 1966

U. Ingard, "Wechselwirkung Zwischen Schall und Strömung" (invited)

Solid State Devices Research Conference, Chicago, Illinois
June 15-17, 1966

A. H. Mc Whorter and H. A. Haus, Amplification at Subcritical Drift Velocities

Second Conference on Coherence and Quantum Optics, Rochester, New York
June 22-24, 1966

C. Freed and H. A. Haus, Intensity Fluctuations of Gas Lasers below and above Threshold of Oscillations (invited)

Lectures on Active Plasma Effects in Solids, Technische Universität Berlin, Fakultat für Electrotechnik, Institut für Hochfrequenztechnik, Berlin, West Germany
June 27, 30, July 4, 7, 11, 14, 1966

A. Bers, Plasma Waves in Solids (invited)

Electron-Phonon Interactions (invited)

Transit-Time Effects and Devices (invited)

Twenty-fourth Annual Conference on Electron Device Research, California Institute of Technology, Pasadena, California

June 29-July 1, 1966

> M. A. Lieberman, Electron-Beam Excitation of Ion Oscillations in an Electron Cyclotron Resonance Discharge

Fifth International Symposium on Rarefied Gas Dymanics, Oxford, England

July 4-7, 1966

> R. M. Logan, J. C. Keck, and R. E. Stickney, Simple Classical Model for the Scattering of Gas Atoms from a Solid Surface II. Additional Analyses and Comparisons

International Symposium on MHD Electrical Power Generation (sponsored by the European Nuclear Energy Agency and the International Atomic Energy Agency), Salzburg, Austria

July 4-8, 1966

> G. A. Brown and E. K. Levy, Liquid-Metal Magnetic Hydrodynamic Power Generation with Condensing Ejector Cycles

> W. D. Jackson and E. S. Pierson, Material Limitations in the MHD Induction Generator

> J. L. Kerrebrock, M. A. Hoffman, and G. C. Oates, A Large Nonequilibrium MHD Generator

> C. A. Mc Nary and W. D. Jackson, Brayton Cycle Magnetohydrodynamic Power Generation

Pattern Recognition Course Lecture, Massachusetts Institute of Technology, Cambridge, Massachusetts

July 7, 1966

> W. M. Siebert, Stimulus Transformations in the Peripheral Auditory System (invited)

Symposium on Aerospace Measurement Techniques, Electronics Research Center, National Aeronautics and Space Administration., Cambridge, Massachusetts

July 7-8, 1966

> G. Fiocco, The Measurement of Atmospheric Parameters by Laser (invited)

> M. W. P. Strandberg, Spectroscopy at Microwave Frequencies (invited)

Physikalische Gesellschaft zu Berlin, Berlin, West Germany

July 8, 1966

> A. Bers, Stability Theory for Plasmas and Dispersive Media (invited)

Technische Universität, Berlin, West Germany

July 16-25, 1966

> W. A. Rosenblith, Lectures (invited)

Lecture, Heinrich Herz Institut für Schwingungsforschung, Berlin, West Germany
July 17, 1966

U. Ingard, Some Problems in Plasma and Laser Acoustics (invited)

122nd Meeting, American Astronomical Society, Ithaca, New York
July 25-28, 1966

R. J. Allen and A. H. Barrett, Absolute Measurements of the Radio Flux from Cassiopeia A and Taurus A at 3.64 and 1.94 cm

R. J. Allen and A. H. Barrett, Secular Variations in the Centimeter Radio Flux from the Quasi-stellar Sources 3C273 and 3C279

J. M. Moran, Jr. and D. H. Staelin, Observations of the Moon near One Centimeter Wavelength

A. E. E. Rogers and A. H. Barrett, Radio Detection of Interstellar $O^{18}H^1$

D. H. Staelin and R. W. Neal, Observations of Venus and Jupiter near 1-cm Wavelength

Conference on Data Acquisition and Processing in Biology and Medicine, Rochester, New York
July 27, 1966

R. Suzuki, Simulation Study of a Hand Controlled by Myoelectric Signals

XVIIIth International Congress of Psychology, Moscow, U.S.S.R.
August 1-7, 1966

P. A. Kolers, Experimental Analysis of Reading

K. N. Stevens, On the Relations between Speech Movements and Speech Perception

Gordon Research Conference on Plasma Physics, Crystal Mountain, Washington
August 8-12, 1966

A. Bers, Wave-type Instabilities of Plasmas in Solids (invited)

International Congress of Mathematicians, Moscow, U.S.S.R.
August 14-28, 1966

P. Elias, Networks of Gaussian Channels with Applications to Feedback Systems (invited)

Western Electric Show and Convention, Los Angeles, California
August 23-26, 1966

J. Andersen and H. B. Lee, Realizability Conditions for the Driving-Point Impedances of RLCT Networks in Which the Reactances Exhibit Semiuniform Loss

A. B. Baggeroer, Maximum a posteriori Interval Estimation

D. L. Snyder, A Theory of Continuous Nonlinear Recursive Filtering with Application to Optimum Analog Demodulation

H. L. Van Trees, Detection and Continuous Estimation Theory

International Astronomical Union Symposium No. 31, "Radio Astronomy and the Galactic System," Leiden, Netherlands

August 25-31, 1966

> B. F. Burke, Comparison of the Distribution and Motion of Atomic Hydrogen in Our Own Galaxy with Other Spiral Systems (invited)

American Physical Society Meeting, Mexico City, D.F.

August 29-31, 1966

> W. P. Allis and J. C. de Almeida Azevedo, The Coupling of Waves across Hybrid Resonance in Inhomogeneous Plasmas

Solid State Theory Seminar, Hughes Research Laboratories, Malibu, California

August 31, 1966

> J. W. Gadzuk, Self-Consistent Field Theory of a Point Impurity in the Surface Region of a Bounded Electron Gas (invited)

JOURNAL PAPERS ACCEPTED FOR PUBLICATION

> (Reprints, if available, may be obtained from the Document Room, 26-327, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.)

R. J. Allen and A. H. Barrett, Absolute Measurements of the Radio Flux from Cassiopeia A and Taurus A at 3.64 and 1.94 cm (Abstract) (Astron. J.)

R. J. Allen and A. H. Barrett, Secular Variations in the Flux from Radio Sources at 3.64 and 1.94 cm (Abstract) (Astron. J.)

J. Andersen and H. B. Lee, Network Synthesis Using Lossy Reactances (IEEE Trans. (CT))

A. H. Barrett, J. W. Kuiper and W. B. Lenoir, Balloon-borne Observations of Microwave Emission by Molecular Oxygen in the Terrestrial Atmosphere (J. Geophys. Res.)

S. Bromberger, Questions (J. Philos.)

S. C. Brown, A Bibliographic Search by Computer (Phys. Today)

P. Chorney, A. Bers, and P. Penfield, Jr., Further Generalization of Waveguide Theorems (IEEE Trans. (MIT))

T. H. Dupree, A Perturbation Theory for Strong Plasma Turbulence (Phys. Fluids)

C. Freed and H. A. Haus, Photoelectron Statistics Produced by a Laser Operating below and above the Threshold of Oscillation (IEEE J. Quantum Electronics)

A. Fukumoto and M. W. P. Strandberg, Fermi Surface in Gallium Determined from the Radio-Frequency Size Effect (Phys. Rev.)

J. W. Gadzuk, Single-Phonon Accommodation Coefficients (Phys. Rev.)

J. W. Gadzuk, Theory of Atom-Metal Interactions I. Alkali Atom Adsorption (Surface Sci.)

J. W. Gadzuk, Theory of Atom-Metal Interactions II. One-Electron Transition Matrix Elements (Surface Sci.)

JOURNAL PAPERS ACCEPTED FOR PUBLICATION (continued)

D. B. Geselowitz, On Bioelectric Potentials in an Inhomogeneous Volume Conductor (Biophys. J.)

E. B. Hooper, Jr. and G. Bekefi, A Laser Interferometer for Repetitively Pulsed Plasmas (J. Appl. Phys.)

K. C. Koerber, R. R. Pfeiffer, W. B. Warr, and N. Y. S. Kiang. Spontaneous Spike Discharges from Single Units in the Cochlear Nucleus after Destruction of the Cochlea (Exptl. Neurol.)

G. H. Matthews, A Conjecture Concerning Non Context-Free Languages (ICC Bulletin)

L. M. Mendell, Physiological Properties of Unmyelinated Fiber Projection to the Spinal Cord (Exptl. Neurol.)

J. M. Moran, Jr. and D. H. Staelin, Observations of the Moon near 1-cm Wavelength (Abstract) (Astron. J.)

P. Penfield, Jr., Thermodynamics and the Manley-Rowe Equations (J. Appl. Phys.)

C. H. Perry, D. P. Athans, E. F. Young, J. R. During, and B. R. Mitchell, Far Infrared Spectra of Palladium Compounds III. Tetrahalo, Tetraammine, and Dihalodiammine Complexes of Palladium (II)

A. E. E. Rogers and A. H. Barrett, Radio Detection of Stellar OH (Abstract) (Astron. J.)

J. G. Siambis and T. G. Northrop, Magnetic Field Geometry and the Adiabatic Invariants of Particle Motion (Phys. Fluids)

D. H. Staelin and R. W. Neal, Observations of Venus and Jupiter near 1-cm Wavelength (Abstract) (Astron. J.)

B. E. Warren and R. L. Mozzi, Multiple Scattering of X-rays by Amorphous Samples (Acta Cryst.)

J. F. Waymouth, Perturbation of Electron Energy Distribution by a Probe (J. Appl. Phys.)

T. F. Weiss, A Model of the Peripheral Auditory System (Kybernetik)

M. L. Wiederhold and W. T. Peake, Efferent Inhibition of Auditory Nerve Responses: Dependence on Acoustic Stimulus Parameters (Science)

D. R. Wilkins and E. P. Gyftopoulos, Transport Phenomena in Low-Energy Plasmas (J. Appl. Phys.)

D. R. Wilkins and E. P. Gyftopoulos, Thermionic Converters Operating in the Ignited Mode. Part I. Theoretical Output Current Characteristics (J. Appl. Phys.)

D. R. Wilkins and E. P. Gyftopoulos, Thermionic Converters Operating in the Ignited Mode. Part II. A Quasi-Equilibrium Model for the Interelectrode Plasma (J. Appl. Phys.)

S. H. Zisk, Brightness Distributions of Radio Sources at 2-cm Wavelength (Science)

LETTERS TO THE EDITOR ACCEPTED FOR PUBLICATION

A. Fukumoto and M. W. P. Strandberg, Radio-Frequency Size-Effect Line Shape for Gallium (Phys. Letters)

W. D. Johnston, Jr. and J. G. King, Measurement of Velocity Distributions of Atoms Evaporating from Liquid Helium II (Phys. Rev. Letters)

H. B. Lee and P. J. Murphy, On the Natural Frequency Realizing Ability of Two Element Kind Structures (IEEE Trans. (CT))

LETTERS TO THE EDITOR ACCEPTED FOR PUBLICATION (continued)

P. Penfield, Jr., Electromagnetic Momentum in a Material Medium (Proc. Inst. Elec. Engrs. (London))

P. Penfield, Jr., Electromagnetic Momentum Associated with Waveguide Modes (Proc. Inst. Elec. Engrs. (London))

P. Penfield, Jr., Unresolved Paradox in Circuit Theory (Proc. IEEE)

A. E. E. Rogers, J. M. Morgan, P. O. Crowther, B. F. Burke, M. L. Meeks, J. A. Hall, and G. M. Hyde, Interferometric Study of Cosmic Line Emission at OH Frequencies (Phys. Rev. Letters)

E. Rojas, J. Y. Lettvin, and W. F. Pickard, A Demonstration of Ion-Exchange Phenomena in Phospholipid Mono-molecular Films (Nature)

H. M. Schneider, On Linearization of the Oscillations in a Plasma Slab (Phys. Fluids)

TECHNICAL REPORTS PUBLISHED

(These and previously published technical reports, if available, may be obtained from the Document Room, 26-327, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.)

442   Thomas J. Fessenden, Pulsed Electron-Cyclotron Resonance Discharge Experiment

446   Ralph Alter, Bioelectric Control of Prostheses

SPECIAL PUBLICATIONS

R. Alter, Special Aspects of Prosthetic Devices (Proc. Second Conference on Information and Control Process in Living Systems, Civic Palisades, California February 27-March 2, 1966

S. Bromberger, Why-Questions (in Mind and Cosmos, Vol. III in University of Pittsburgh Series in the Philosophy of Science, University of Pittsburgh Press, 1966), pp. 86-111

T. J. Fessenden and L. D. Smullin, A Simple Technique for Measuring the Density of a Plasma in a Metal Box (Proc. Seventh International Conference on Phenomena in Ionized Gases, edited by B. Perović and D. Tošić, Institute of Nuclear Sciences, Beograd, July 1966)

T. J. Fessenden and L. D. Smullin, Electron Production and Loss in a Plasma Produced by Pulsed Microwaves at the Electron-Cyclotron Resonant Frequency (Proc. Seventh International Conference on Phenomena in Ionized Gases, edited by B. Perović and D. Tošić, Institute of Nuclear Sciences, Beograd, July 1966), pp. 201-204

INTRODUCTION

This report, the eighty-third in a series of quarterly progress reports issued by the Research Laboratory of Electronics, contains a review of the research activities of the Laboratory for the three-month period ending August 31, 1966. Since this is a report on work in progress, some of the results may not be final.

GENERAL PHYSICS

I. MICROWAVE SPECTROSCOPY[*]

Academic and Research Staff

Prof. M. W. P. Strandberg
Prof. R. L. Kyhl
Dr. J. M. Andrews

Dr. M. U. Palma
Dr. Maria B. Palma-Vittorelli

A. Fukumoto
J. G. Ingersoll
J. D. Kierstead

Graduate Students

J. U. Free, Jr.
R. M. Langdon, Jr.

M. K. Maul
T. E. McEnally

S. R. Reznek
B. N. Yung

A. ELECTRONIC CONTRIBUTION TO 9-GHz ULTRASONIC ATTENUATION IN THIN FILMS AT LOW TEMPERATURES

The low-temperature ultrasonic reflection coefficients of thin films of tin have been studied in order to determine the electronic contribution to the ultrasonic attenuation coefficient at 9 GHz. In these experiments, tin films between $1.2-\mu$ and $3.6-\mu$ thickness have been deposited onto the end of an x-cut quartz rod. The opposite end of the quartz rod was excited piezoelectrically by means of an X-band re-entrant cavity. Ultrasonic echo pulses reflected from the tin film back to the re-entrant cavity were monitored as the film was switched between the superconducting and normal states by means of an external magnetic field. A diagram of the experimental apparatus is shown in Fig. I-1. The details of the ultrasonic energy-flow within a film, x_0 cm thick, are shown in
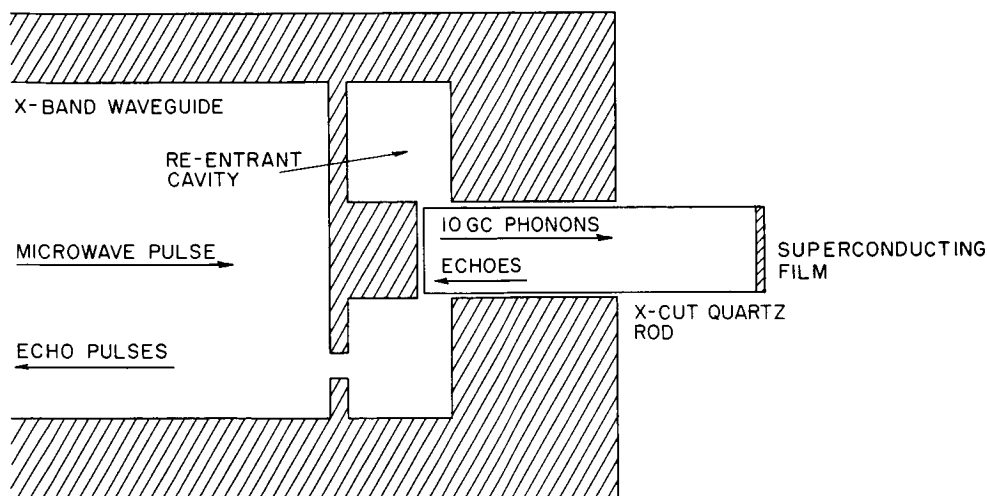
Fig. I-1. Apparatus for the observation of the microwave ultrasonic reflection coefficient of thin superconducting films.

Fig. I-2. The quartz-metal boundary is characterized by a power reflection coefficient R, and a power transmission coefficient T, so that

$$R + T = 1. \tag{1}$$

Losses in the tin film are characterized by an attenuation coefficient a (cm^{-1}). The far surface of the tin film is assumed to be perfectly reflecting, but rough so that only a fraction, F, of the incident power is specularly reflected with no loss in phase coherence. We shall deal only with the case of normal incidence. (The oblique incidence shown in Fig. I-2 is only for clarity in path presentation.)
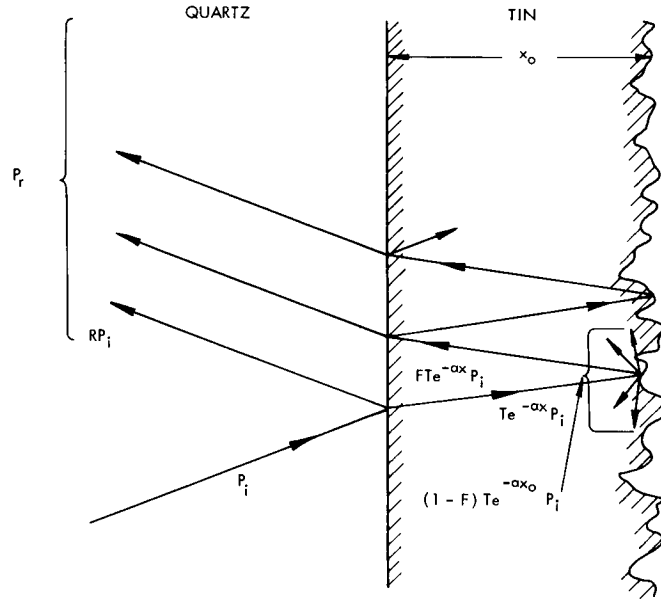
Fig. I-2. Multiple reflection of ultrasound in a thin film.

The total reflected power, P_r (watts), can be obtained as a function of the incident power, P_i (watts), by summing the geometric series representing multiple reflections within the metallic film. The ratio is denoted Γ^2.

$$\Gamma^2 = \frac{R + F\,e^{-2ax_o} - 2\sqrt{FR}\,e^{-ax_o}\cos\phi}{1 + FR\,e^{-2ax_o} - 2\sqrt{FR}\,e^{-ax_o}\cos\phi} \tag{2}$$

The quantity ϕ (radians) is a measure of the phase retardation within the film.

$$\phi = \frac{4\pi x_o}{\lambda} = \frac{4\pi\nu x_o}{v_s}, \tag{3}$$

where λ (cm) is the ultrasonic wavelength within the tin film, ν (sec^{-1}) is the frequency, and v_s (cm/sec) is the ultrasonic velocity.

The power reflection coefficient Γ^2 approaches unity in two limits. When $R \rightarrow 1$, no power enters the metallic film, and it is all returned to the generator (quartz losses are neglected). When $a = 0$ and $F = 1$, no loss or dispersion of ultrasound takes place within the metallic film. Again, in the case of normal incidence, all of the ultrasonic power is returned to the generator.

The power reflection coefficient, R, at the interface between two solids can be represented by

$$R = \left(\frac{Z_1 - Z_2}{Z_1 + Z_2}\right)^2, \tag{4}$$

where, for a lossless medium, $Z_i = \rho_i (v_s)_i$ (gm cm^{-2} sec^{-1}) is the ultrasonic impedance, ρ_i (gm cm^{-3}) is the density of medium i, and $(v_s)_i$ is the ultrasonic velocity in medium i. For a quartz-tin interface, the theoretical power reflection coefficient, evaluated for the longitudinal mode, is 0.048. For a nonideal interface, or when ultrasonic losses are large, the theoretical power reflection coefficient could be much larger.

In the case of small R, we have an approximation of Eq. 2 that becomes exact in the limit $x_o \rightarrow 0$.

$$\Gamma^2 \approx F\, e^{-2 a x_o}. \tag{5}$$

The ultrasonic attenuation coefficient a (cm^{-1}) is viewed as being composed of two portions: (i) a residual attenuation a_o (cm^{-1}) caused by grain-boundary scattering, crystalline imperfections, impurities, and other defects, and (ii) a_{elec} (cm^{-1}) caused by the conduction electrons. The electronic component varies when the metal is switched between the normal and the superconducting states according to

$$\frac{a_s}{a_n} = \frac{2}{e^{2\epsilon(T)/kT} + 1}, \tag{6}$$

where $\epsilon(T)$ (erg) is the superconducting energy gap, k (erg°K^{-1}) is Boltzmann's constant, and T (°K) is the absolute temperature.

In the limit $\left(RF^{-1} e^{2 a x_o}\right) \ll 1$, Eq. 5 is applicable, and the ratio of the ultrasonic reflection coefficients for the superconducting versus the normal state of the metallic film can be expressed by

$$D(T)\ (db) = 20 \log_{10} (\Gamma_s / \Gamma_n) = 20(a_n - a_s)\, x_o \log_{10} e \qquad RF^{-1} e^{2 a x_o} \ll 1. \tag{7}$$

The temperature-dependent $a_s(T)$ can be eliminated by substituting Eq. 6. in Eq. 7.

$$D(T) = 20a_n(T) \tanh \left[\frac{\epsilon(T)}{kT} \right] x_o \log_{10} e. \tag{8}$$

Below 4°K, it is reasonable to assume that $a_n(T) = a_n(0)$ for a thin, polycrystalline film. In this case,

$$D(0) = \frac{D(T)}{\tanh \left[\dfrac{\epsilon(T)}{kT} \right]} = 20a_n(0) \; x_o \log_{10} e. \tag{9}$$

We have made observations of D(T) for thin films of tin with thicknesses in the range 1.2-3.6 μ. These measurements were made at temperatures below 2°K, and the temperature normalization defined by Eq. 9 yields corrections only of the order of a few parts per thousand. The BCS value for the superconducting energy gap was used: $[2\epsilon(0) = 3.51 \, kT_c]$. These normalized measurements are shown in Fig. I-3.

Referring to Eq. 2, it is evident that as $x_o \rightarrow \infty$, $D(0) \rightarrow 0$. At intermediate values of x_o, $D(0)$ should exhibit oscillations caused by $\cos \phi$. The theoretically expected behavior of $D(0)$ is sketched in Fig. I-4. The magnitude of the oscillation is controlled by the
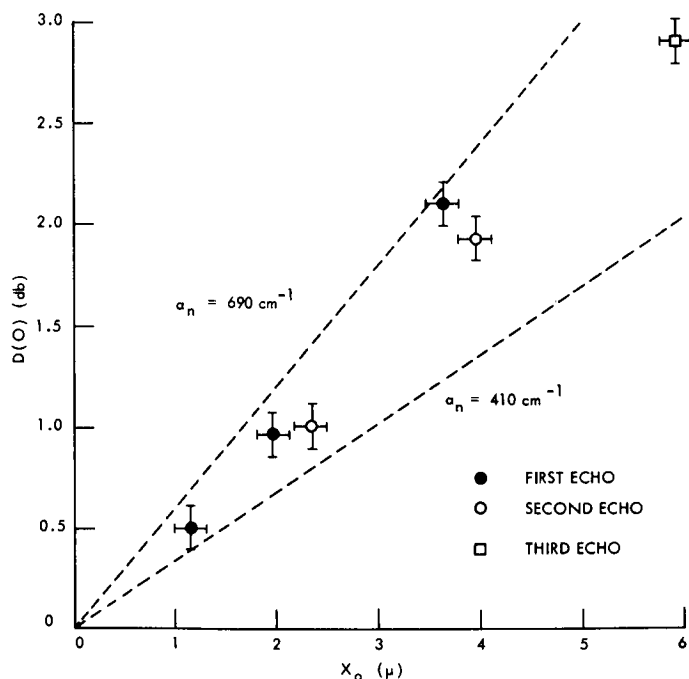
Fig. I-3. Experimental values of the ratio of the microwave ultrasonic reflection coefficient of a tin film as it is switched between the superconducting and normal states.

$$D(0) = 20 \log_{10} \left(\frac{\Gamma_s}{\Gamma_n} \right).$$

$\alpha_o / \alpha_n = 10$

R = 0.05
F = 1
$\alpha_n = 0.05 \, \mu^{-1}$

3

1

0

10

3

Eq. 9

0

1

3

10

D (0) (db)

7

6

5

4

3

2

1

0

X_o (μ)

0   1   2   3   4   5   6   7   8   9   10
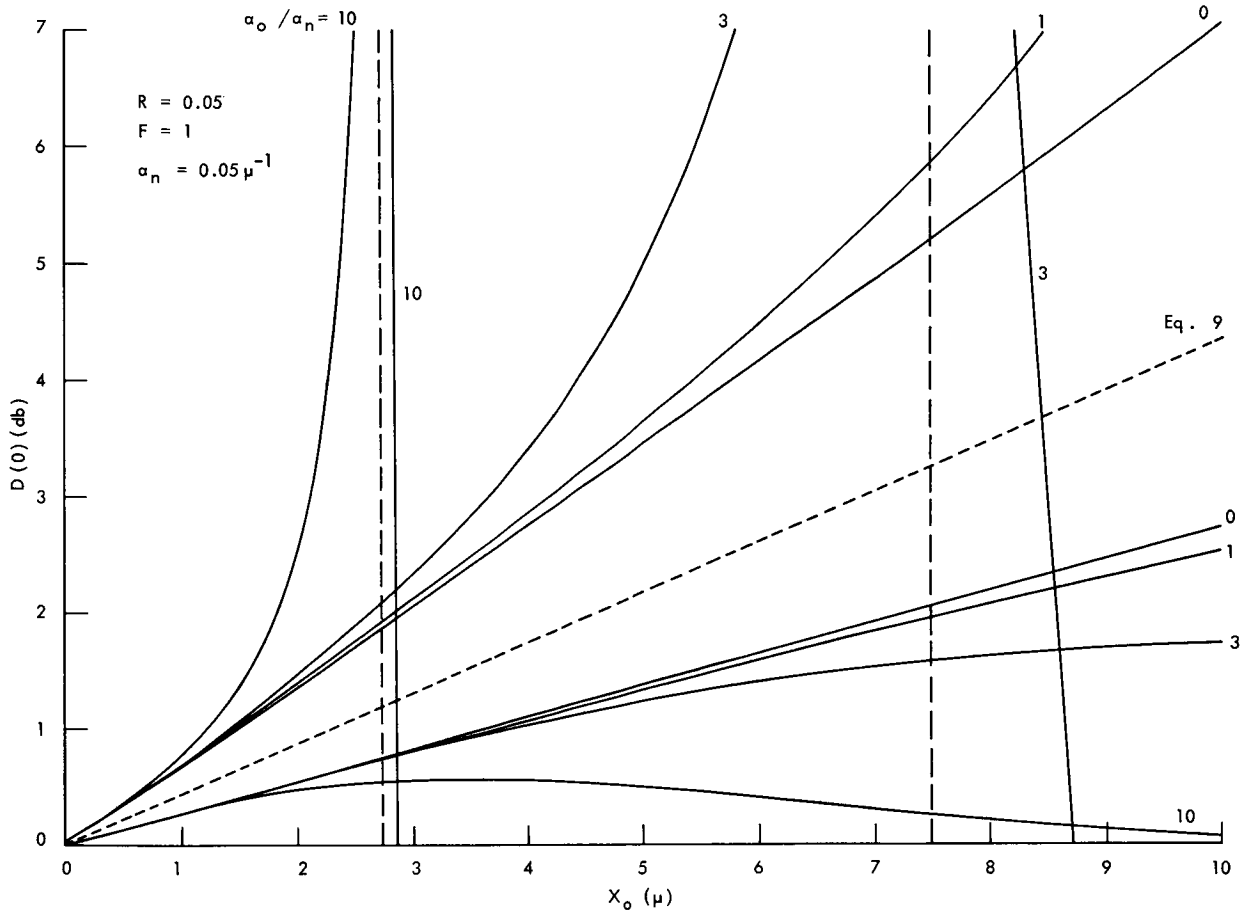
Fig. I-4. Theoretical behavior of $D(0) = 20 \log_{10} (\Gamma_s/\Gamma_n)$, where Γ_s/Γ_n is defined by Eq. 2 with $\alpha = \alpha_o + \alpha_n$. These curves have been drawn for R = 0.05, F = 1, $\alpha_n = 0.05 \, \mu^{-1}$, and $\alpha_o/\alpha_n = 0, 1, 3, 10$. The periodic quantity $\cos \phi$ oscillates rapidly between the two limiting curves shown for each value of α_o/α_n. The two vertical dashed lines show singularities in Γ_s/Γ_n, where no signal can be detected in the normal state. This is caused by complete cancellation of the ultrasonic waves reflected from each of the two surfaces of the metallic film. At values of x_o slightly higher than those that yield a singularity, the fraction Γ_s/Γ_n becomes less than unity. The central dotted line, between the two limiting curves for $\alpha_o/\alpha_n = 0$, represents the limiting behavior for R → 0 (see Eq. 9).

values of R, F, and a_o. If we evaluate the phase retardation ϕ from Eq. 3 for the longitudinal mode in polycrystalline tin ($v_s = 3.3 \times 10^5$ cm sec^{-1}) at 9 GHz, we find that one period of oscillation corresponds to a change in thickness of the tin film equal to 0.18 μ. This value is of the same order of magnitude as our measurement errors, so, without an appreciable increase in our optical resolution, it will not be possible to observe the oscillations in our experimental data.

In Fig. I-3 there is good agreement among the data taken from multiple echoes. The small discrepancy in the data for different film thicknesses could be interpreted as being indicative of the magnitude of the oscillations. Further experiments, however, are needed to establish whether or not this is actually the case.

A theoretical expression for the electronic contribution to the ultrasonic attenuation coefficient in metals has been worked out by Pippard.[1] It is given by

$$a_n = \left[\frac{\pi^2}{6} \left(\frac{3\rho}{\pi} \right)^{1/3} \left(\frac{N_o}{W} \right)^{4/3} \frac{h}{v_s^2} \right] \nu n^{4/3} \; cm^{-1}, \tag{10}$$

where ρ (gm/cm^3) is the density of the metal, N_o (atoms/mole) is Avogadro's number, W (gms/mole) is the atomic weight of the metal, h (erg sec) is Planck's constant, v_s (cm/sec) is the ultrasonic velocity in the metal, ν (sec^{-1}) is the frequency of the ultrasound, and n is the effective number of free electrons per atom.

The chemical valance of tin can be either 2 or 4. Equation 10, evaluated for the longitudinal ultrasonic mode in tin at 9 GHz, yields

$$\text{Theoretical} \begin{cases} a_n = 380 \; cm^{-1} & \text{2 free electrons/atom} \\ a_n = 650 \; cm^{-1} & \text{3 free electrons/atom} \\ a_n = 950 \; cm^{-1} & \text{4 free electrons/atom} \end{cases}$$

The two limiting slopes of our experimental data shown in Fig. I-3 yield

$$\text{Experimental} \begin{cases} a_n = 690 \; cm^{-1} \\ a_n = 410 \; cm^{-1} \end{cases}$$

These values suggest that some number between 2 and 3 electrons per atom in metallic tin is effective in the electronic contribution to ultrasonic attenuation.

The author would like to acknowledge the experimental assistance of John U. Free, Jr.

J. M. Andrews, Jr.

References

1. A. B. Pippard, Rept. Progr. Phys. <u>23</u>, 176 (1960), Eq. 68, p. 242.

Academic and Research Staff

N67 14623

Prof. A. H. Barrett	Prof. L. B. Lenoir	Dr. S. H. Zisk
Prof. B. F. Burke	Prof. D. H. Staelin	Patricia P. Crowther
Prof. M. Loewenthal		E. R. Jensen

Graduate Students

R. J. Allen	G. D. Papadopoulos	A. E. E. Rogers
N. E. Gaut	E. C. Reifenstein III	K. D. Thompson
J. M. Moran, Jr.		T. L. Wilson

A. OH INTERFEROMETRY

In order to investigate further the physical conditions responsible for the anomalously intense sources of 18-cm line emission at radio frequencies, an interferometric study was undertaken jointly with M. L. Meeks and G. M. Hyde of Lincoln Laboratory, M.I.T., to determine the angular dimensions of the emitting regions. The interferometer was composed of the Millstone (84-ft) and Haystack (120-ft) antennas of Lincoln Laboratory, with a baseline of approximately 3800 λ at 18 cm, along a line nearly 20° east of north. Most of the observations were made with both antennas circularly polarized in the same sense.

The signals from the two antennas were effectively crosscorrelated by a phase-switching scheme. The sum and difference of the IF outputs from the two receivers were autocorrelated, and the difference between these autocorrelation functions was taken. A common local-oscillator signal was derived from reference signals carried along a transmission line that was servo-controlled to maintain constant electrical length. An IF delay for white-fringe compensation was unnecessary, since the delays were reconstructed in the autocorrelator.

Fringe amplitude and phase information, as a function of frequency, was extracted from the autocorrelation functions by means of a least-squares-fit technique executed by a digital computer. After calibration of the baseline parameters with continuum radio sources of small diameters and known positions, the positions of the emission lines were obtained from fringe phase as a function of hour angle.

The observations have been concentrated, thus far, on the emission regions near the continuum radio sources W3 (IC 1795), W49, Sgr A, and NGC 6334. Table II-1 shows, for the five strongest lines at 1665 MHz in W3, the size limits (under the assumption of a uniformly bright circular disc) of the emitting source derived from the observed fringe amplitude. The uncertainty in fringe amplitude represents the peak observed deviation

Table II-1. Angular sizes and separations of emission
features observed adjacent to W3.

Line Velocity[a] Km/s	Polarization (I.R.E. Convention)	Fringe Amplitude	Effective Source Diameter	Separation from −45.1 Km/s Line
−45.1	Right	1.01 ± 0.05	<15"	
−43.7	Right	1.0 ± 0.1	<20"	<3"
−41.7	Right	1.0 ± 0.2	<25"	<3"
−45.4	Left	1.0 ± 0.1	<20"	<3"
−46.4	Left	1.0 ± 0.1	<20"	<3"

[a]Velocity relative to the local standard of rest under the assumption that rest frequencies are those of the $2\Pi_{3/2}$, J = 3/2, Λ-doublet of $O^{16}H^{1}$.

for 15-minute integration intervals over all local hour angles. All lines had nearly the same phase which allows us to put an upper limit on the angular separation of the individual lines. No significant resolution of the emitting source could be detected, and all of the radiation appears to originate from the same region. More limited observations of the 1667-MHz lines gave the same position for the lines at −42.2 Km/sec and −44.7 Km/sec within 10 seconds of arc, subject to a possible lobe ambiguity owing to the small local-hour-angle coverage of the observations at this frequency.

Since all hour angles were covered at 1665 Mhz, the position was unambiguously determined from the observations made June 7th through June 19th (Epoch 1950.0) to be

$$\alpha = 02^{h} 23^{m} 14.3^{s} \pm 1.5^{s}$$

$$\delta = 61° 38' 57" \pm 10".$$

The dominant contribution to these errors arises from uncertainty in the derived interferometer baseline. The rms fluctuations (for a 15-minute integration) because of noise in the observations of the strongest line (−45.1 Km/sec) was 3 per cent in fringe amplitude and 3° in fringe phase.

A search was made in the Palomar Observatory Sky Atlas, for possible optical identifications. The observed position falls just within the boundary of the nebulosity, midway between two faint stars, neither of which is within the position uncertainty.

The source of the line emission is clearly of an unusual nature. For example, the

observed angular size limit implies a brightness temperature of at least 2×10^6 °K for the line at -45.1 Km/sec. The apparent linear dimension of the source, under the assumption that it is located at the distance of W3 (1700 parsec), is less than 0.1 parsec. Sgr A is also a single point source, less than 20" in size. W49 and NGC 6334 each appear to be slightly more complex, and are apparently each double. The individual components are, however, unresolved by the interferometer.

In summary, all OH emission sources observed thus far are of very small angular size, although more than one point source can be associated with a given H II region. Each point source usually has more than one velocity component associated with it.

B. F. Burke, J. M. Moran, A. E. E. Rogers

B. LOW-FREQUENCY APERTURE SYNTHESIS OF DISCRETE RADIO SOURCES

A program of observations has just been completed at the National Radio Astronomy Observatory, Green Bank, West Virginia. We measured the fringe visibilities of approximately 24 of the brightest radio sources, at a frequency of 234 MHz, using the long-baseline interferometer with additional equipment constructed at M.I.T.

From the results of these observations, the structures of many of these sources are being obtained by a variety of digital-computer processings, including smoothing, interpolation, inverse Fourier transformation, and model fitting.

The NRAO interferometer consists of two radio telescopes, of 85-ft diameter, capable of operating at any of 6 baselines from 1200 m to 2700 m in separation, along an azimuth of ~50° (H = $4^h 50^m$, D = -22° 07' approximately). The receiver is of the double-sideband variety, with an intermediate frequency band of 2-12 MHz (Read, 1963). We fitted the telescopes with 234-MHz feeds and front-end electronics, which we then led into the NRAO IF delay and correlation system. The first stage in the processing of the data was finished at NRAO, and the remaining work will be done here during the next 6-8 months.

We selected, for the most part, small extragalactic sources for this program in an attempt to detect larger-scale diffuse halos around the major components that might be depositories for old, lower-energy electrons or the residues of events that preceded those responsible for the major components of the radio sources observed at present. Our ultimate angular resolution (for those sources for which maps can be obtained) will be of the order of 1 min (arc), compared with the 2-5 min (arc) sizes of the major components. In some cases the primary beam of the telescopes (diameter = 3° (arc)) included other comparable sources besides the one of interest, and the large minimum spacing of the interferometer (1200 m = 960 wavelengths) will not allow the several sources to be separated (a problem analogous to the multiple responses of a grating array). It is hoped that next year, when a third telescope will have been constructed at NRAO that

will afford a spacing as small as 100 m (80 wavelengths), the missing region of the fringe-visibility plane can be filled in and these complicated areas can be mapped completely. In the meantime, we are analyzing those sources whose structure can now be determined with some certainty.

S. H. Zisk

C. SPECTRUM MEASUREMENTS OF VENUS AND JUPITER

Microwave spectroscopy can provide information about the identity, abundance, and distribution of any atmospheric constituents which have appropriate microwave reso-nances. As part of a search for atmospheric spectral features, the planets Venus and Jupiter were observed from January to March, 1966, at wavelengths near 1 centimeter.

The observations were made with the 28-ft antenna[1] at Lincoln Laboratory, M.I.T.; and a five-channel microwave radiometer[2] operating simultaneously at 19.0, 21.0, 22.235, 23.5 and 25.5 GHz. The system sensitivity evaluated at the antenna feed, and including baseline drift, was 2.5°K for a 1-sec time constant.

The planetary observations were calibrated by comparison with lunar observations.
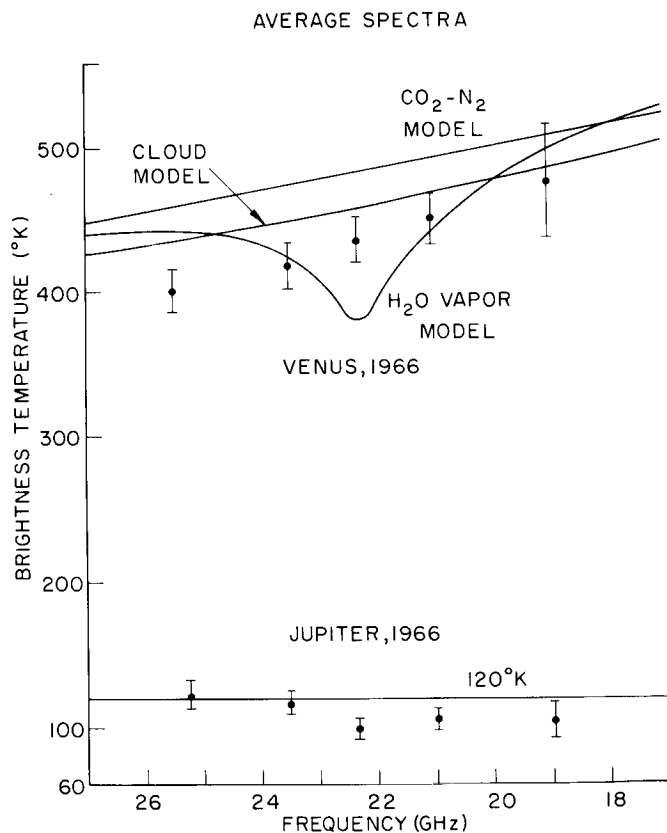
Fig. II-1. Observed and theoretical average spectra of Venus and Jupiter.

The comparison was made utilizing the measured antenna patterns, as described ear-
lier[3,4]. The measurements were corrected for atmospheric absorption with an accuracy
of approximately 1 or 2 per cent.[3] The atmospheric absorption was determined by
means of a concurrent series of solar-extinction atmospheric absorption measurements
and measurements of ground-level humidity.

Venus was observed on 13 days between January 11 and March 15, 1966. The average
spectrum is shown in Fig. II-1 and is listed in Table II-2. The error brackets represent
the relative rms accuracies of the measurements. The results tabulated in Table II-2
include the absolute accuracies, which incorporate uncertainties in the antenna pointing,
in the measured antenna patterns, and in the assumed lunar brightness temperatures.

The observed spectrum can be compared with the theoretical spectra shown in
Fig. II-1. The illustrated water vapor spectrum corresponds to an N_2-CO_2-H_2O model
atmosphere with 50 atm surface pressure and 650°K surface temperature. The water-
vapor density was assumed to be 60 gm/m^3 at the surface and to have a constant mixing
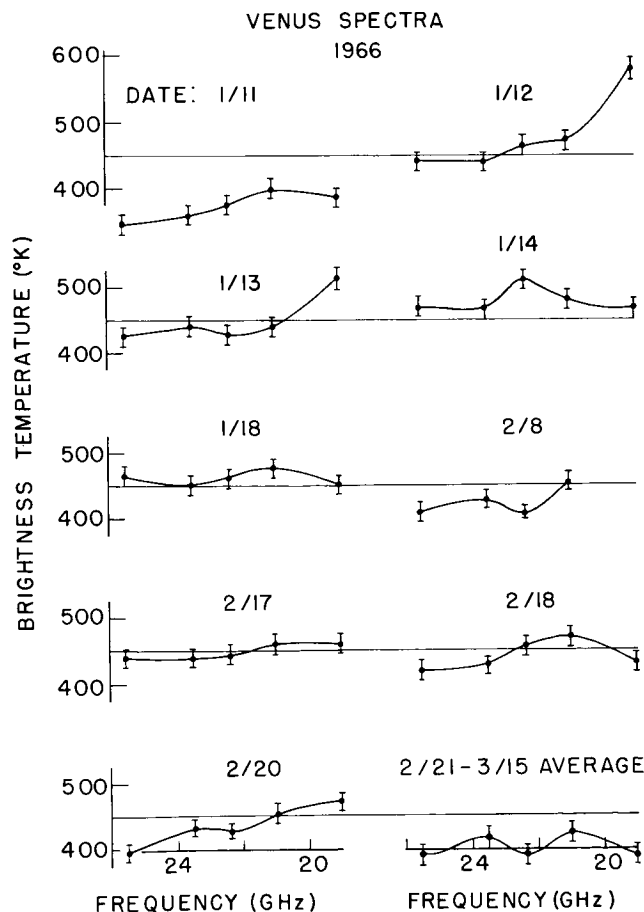
Fig. II-2. Observed spectrum of Venus as a function of date.

Table II-2. Results of spectrum measurements.

Frequency	Venus	RMS Error		Jupiter	RMS Error	
	$\overline{T_B}$	Rel.	Abs.	$\overline{T_B}$	Rel.	Abs.
(GHz)	(°K)	(%)	(%)	(°K)	(%)	(%)
19.0	477	±8	±12	105	±20	±20
21.0	451	±4	±9	106	±9	±10
22.235	436	±4	±9	98	±10	±11
23.5	418	±4	±9	116	±8	±9
25.5	400	±4	±9	123	±7	±9

JUPITER SPECTRA
1966
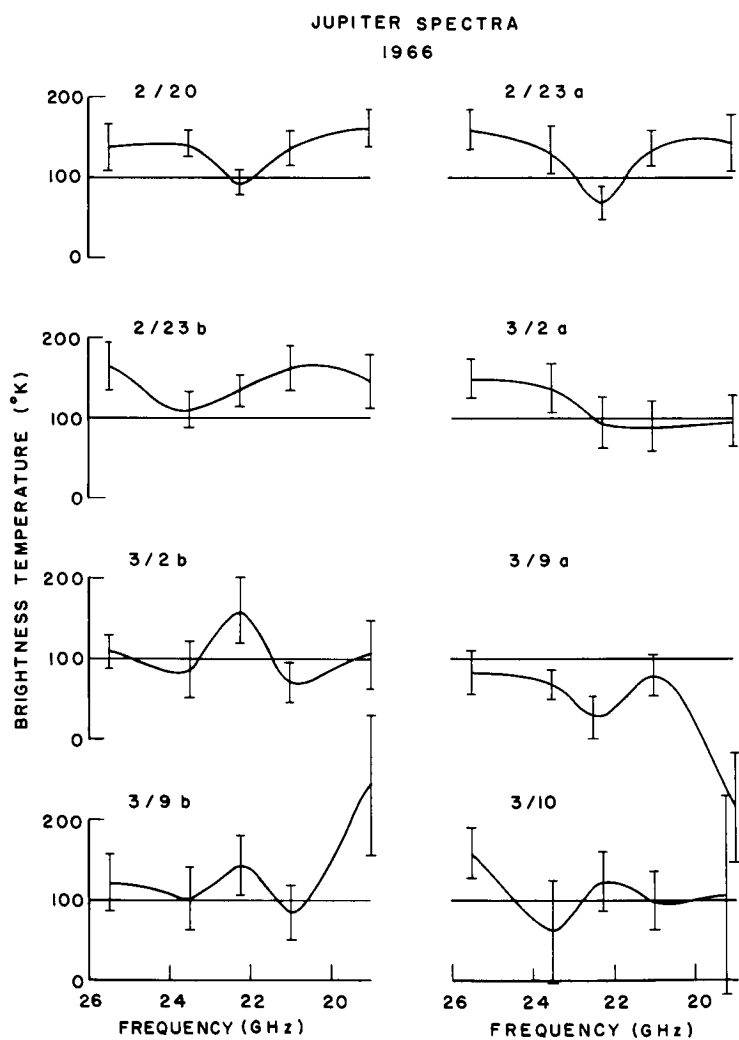
Fig. II-3. Observed spectrum of Jupiter as a function of date.

ratio in all regions hotter than 273°K. This mixing ratio corresponds to a typical winter day. Both of the nonresonant model spectra are consistent with the observed spectrum and the absolute error brackets.

The observed spectrum of Venus is shown as a function of date in Fig. II-2. The error brackets correspond to receiver noise only. There are no apparent fluctuations in spectral shape which are not consistent with either the illustrated error brackets or with instrumental effects such as antenna-pointing errors.

Jupiter was observed on 7 days, yielding the average spectrum shown in Fig. II-1 and tabulated in Table II-2. The observed average spectrum is consistent with a flat spectrum near 110°K. The observed spectra are plotted as a function of date in Fig. II-3; the fluctuations appear greater than the fluctuation of Venus. Although this effect could be instrumental, it could also originate from the very rapid rotation of Jupiter, one revolution every 9.9 hours. The sub-Earth point of Jupiter is plotted as a function of time in Fig. II-4, using System II coordinates. Comparison of Figs. II-3 and II-4 shows that the three
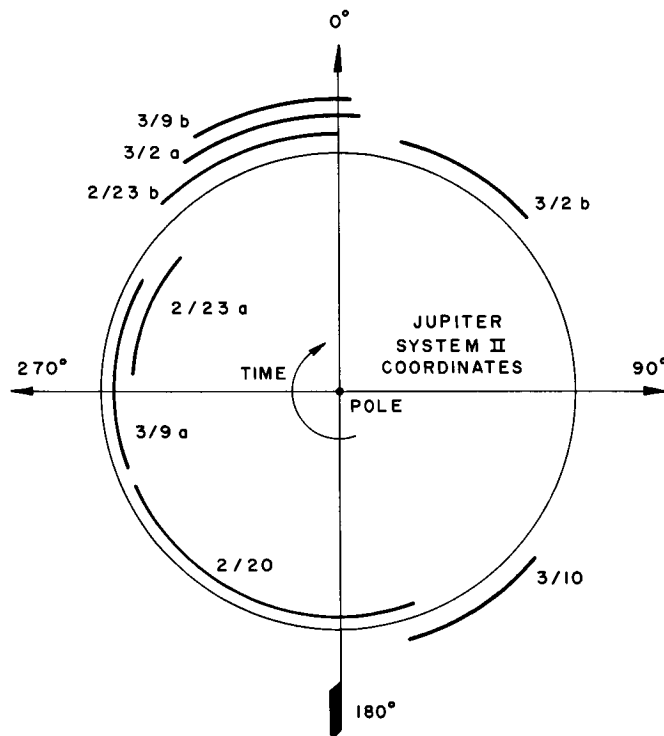
Fig. II-4. System II Longitude of the Jovian sub-Earth point.

spectra exhibiting dips near 22.235 GHz occurred on the same side of Jupiter, and the remainder occurred on the other side. Future observations would be

necessary to confirm this Jovian meterological effect. Such an effect could be due to NH_3, which is known to be present on Jupiter. NH_3 has many strong molecular inversion lines in this spectral region, including the $J = 3$, $K = 1$ line at 22.2345 GHz.

<div align="right">D. H. Staelin, R. W. Neal, S. E. Law, E. C. Reifenstein</div>

References

1. D. H. Staelin, Quarterly Progress Report No. 69, Research Laboratory of Electronics, M.I.T., April 15, 1963, pp. 23-25.

2. D. H. Staelin and A. H. Barrett, Quarterly Progress Report No. 78, Research Laboratory of Electronics, M.I.T., July 15, 1965, pp. 21-24.

3. D. H. Staelin, Sc.D. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, January 1965.

4. D. H. Staelin and A. H. Barrett, Astrophys. J. 144, 352 (1966).

III. SOLID-STATE MICROWAVE ELECTRONICS*

Academic and Research Staff

Prof. R. P. Rafuse
Dr. D. H. Steinbrecker

Graduate Students

A. M. Davis
Z. Vugrinec

A. STATUS OF RESEARCH

Several new developments in varactor multipliers and crystal mixers will be summarized in this report; details are now available in theses and published journal papers, and other papers are being prepared.

A new theory for high-order varactor multipliers utilizing "step-recovery" diodes has been developed.[1,2,3] The implications of the theory, and our experimental verification thereof, have been somewhat at variance with other theoretical treatments; so much so, in fact, that we have chosen to call such varactors "punch-through" diodes, since this title is more representative of their physical properties.

As a result of the present device technology and the theoretical knowledge gained by our group, we have initiated a program to develop a varactor multiplier from 225 MHz to 2.25 GHz supplying a power output of 40-60 watts. We are attempting to build 2.25-GHz solid-state power sources with DC-to-RF conversion efficiencies in excess of 50%, utilizing available transistor technology and the varactor multiplier discussed here.

A very interesting breakthrough in crystal mixer technology has occurred in our laboratory. The details are available in an S.M. thesis[4] and will not be covered here. Suffice it to say, over-all, single-sideband system noise figures of 2.1 db at 2.7 GHz have been obtained with experimental Texas Instruments GaAs Shottky-barrier diodes and 2.3 db in the same mixer with Sylvania IN416G diodes. At present, a double-sideband noise figure of 2.0 db at 9.6 GHz has been obtained with Sylvania IN415G diodes, which indicates that a single-sideband noise figure of 3 db is attainable by using techniques analogous to those employed in the 2.7-GHz mixer. It should be noted that these noise figures include a 0.78 db IF noise figure and a correction of the gas-tube noise sources to 15.75 db (calibrated against thermal sources). Therefore, if the gas-tube correction is ignored and a standard 1.5 db noise figure is assumed for the IF preamplifier, the noise figures quoted will be approximately 0.22 db optimistic compared with industrial figures.

Work continues in both these areas, and extensions into the millimeter-wave region

are planned. We also plan to begin to make our own Shottky-barrier diodes for the higher frequency mixers (above 30 GHz). An examination of phase-lock properties of avalanche-diode oscillators at 15 GHz continues and will consider injection locking of pulsed modes of operation.

R. P. Rafuse, D. H. Steinbrecker

References

1. R. P. Rafuse, "Recent Developments in Parametric Multipliers," Proc. of the National Electronics Conference, October 1963, pp. 461-470.

2. R. P. Rafuse and D. H. Steinbrecker, "Harmonic Multiplication with Punch-Through Varactors," International Solid-State Circuits Conference, Digest of Technical Papers, Vol. IX, pp. 68-69, February 1966.

3. D. H. Steinbrecker, "Limitations on Parametric Amplifiers and Improved Efficiency Varactor Multipliers," Ph. D. Thesis, M. I. T., July 15, 1966.

4. J. Vugrinec, "The Effect of Image-Frequency Termination in Microwave Mixers Using Shottky-Barrier Diodes," S. M. Thesis, M. I. T., September 2, 1966.

Academic and Research Staff

Professor C. H. Perry

Graduate Students

| Jeanne H. Fertel | J. F. Parrish | N. Tornberg |
| D. Muehlner | J. Reintjes | E. F. Young |

A. TRANSMISSION SPECTRA OF (ABF$_3$) TYPE FLUORIDE PEROVSKITES

1. Introduction

Second-order dipole moments resulting from charge deformation during lattice vibrations are adequate to explain the coupling mechanism for multiphonon processes seen in absorption spectra in the infrared.[1] It is also reported that anharmonic terms in the potential energy function contribute to multiphonon processes.[2] Birman[3] performed a full analysis of the space group for diamond and zinc blende crystal structures to determine general selection rules for electronic transitions, vibrational excitations, and mixtures of both. He applied his results to experimental infrared data of multiphonon combination bands for these structures[4] and was able to assign all observed bands to phonon combinations in agreement with earlier assignments.[5]

The full-space group analysis of the cubic perovskite crystal structure for the determination of selection rules for multiphonon processes has not yet been accomplished; however, the observed multiphonon processes seen in the infrared transmission spectra of the cubic perovskites can be assigned to combinations of phonons at critical points on the lattice dispersion curves of the phonon branches by using the general selection rule of Lax and Burstein,[1] which states that if a crystal has a center of symmetry, multiphonon processes can occur only between phonons coming from different branches in the Brillouin zone.

The transmission spectra of thin single crystals (~70μ) of KMgF$_3$, KMnF$_3$, KCoF$_3$, KNiF$_3$, KZnF$_3$, RbMnF$_3$ and K(50% Mg + 50% Ni)F$_3$ have been viewed from 4000 cm^{-1} t0 20 cm^{-1} at 300°K and 85°K with the use of unpolarized radiation. The absorption bands found in this transmission study and side bands found in a reflectance spectra study of the same materials[6] were assigned to multiphonon processes occurring at the edge of the Brillouin zone. A fitting scheme was proposed and the frequencies of the nine phonon branches at the edge of the zone were determined for this list of fluoride perovskites. Ferguson, Guggenheim, and Wood[7] have reported vibrational intervals observed in the electronic absorption spectra study of the Ni^{++} ion in KNiF$_3$. These

vibrational intervals have been successfully assigned to combinations of electronic and vibrational transitions not in agreement with the assignments of Ferguson, Guggenheim, and Wood.

2. Discussion

Multiphonon processes appear as continuous absorption of photons and are classified as summation processes if two or more phonons are created, or difference processes if a phonon is annihilated as well as some being created. As a crystal is cooled, summation bands become more intense and difference bands become less intense and disappear when the temperature is sufficiently low.[6] Energy and momentum must be conserved for the initial and final states of the process.

Table IV-1. Frequencies of absorption bands from transmission data on thin samples.

	$KMgF_3$	$KMnF_3$	$KCoF_3$	$KNiF_3$	$KZnF_3$	$RbMnF_3$	K (50% Mg + 50% Ni) F_3
a.	403	775	350	324	330	333	407
b.	430		810	354	351	600	710
c.	615			363	362	745	875
d.	632			390	640		
e.	667			585	790		
f.	735			630			
g.	762			667			
h.	794			835			
i.	803			*1210			
j.	827						
k.	857						
l.	911						
m.	943						

Temperature 85°K.

All frequencies in cm^{-1}.

*Seen with thick sample only.

The experimental procedure and the plotted transmission spectra of thin single crystals (~70μ thick) of the fluoride perovskites are published elsewhere.[6] The spectra exhibited the three normal vibrations of the allowed infrared active modes as large absorptions with zero transmission at the positions of the peak frequencies in the reflectance spectra. Table IV-1 lists the absorption bands in the infrared transmission spectra of thin single crystals at 85°K and the largest number of absorptions were found

in $KMgF_3$ and $KNiF_3$. The mixed crystal of K (50% Mg + 50% Ni) F_3, on the other hand, had only 3 observable absorptions as compared with the 21 seen in the pure crystals of $KMgF_3$ and $KNiF_3$. The mixing of the B-ion in the cubic ABF_3 perovskite obscures the infrared absorption bands.

The absorption bands found in the transmission study of thin single crystals were temperature-dependent. The absorptions sharpened and shifted to higher frequencies upon cooling to 85°K and were relatively broad, indicating combinations of lattice vibrations. It did not seem reasonable to assign any of the noted absorptions to pure electronic transitions, as these would be normally quite sharp. The positions of the absorption peaks at 300°K could not be accurately located because of the broad shallow nature of the bands.

The transmission spectrum of each crystal studied at frequencies greater than 1200 cm^{-1} showed no noticeable absorptions, the spectra were flat at approximately 90 per cent transmittance, and the transmission at 300°K and 85°K coincided.

The cubic perovskite crystal structure has three triply degenerate infrared active optical modes, one triply degenerate optically inactive mode, and one triply degenerate acoustical mode.[8] Considering the possible polarizations and degeneracies, the cubic perovskite structure has 9 phonon branches in the Brillouin zone. They are described as follows.

2-TO_1 — transverse optical phonon (lowest frequency) i. r. active

2-TO_2 — transverse optical phonon (middle frequency) i. r. active

2-TO_3 — transverse optical phonon (highest frequency) i. r. active

3-O_4 — optical phonon (frequency unknown) not i. r. or Raman active

1-LO_1 — longitudinal optical phonon (lowest frequency) i. r. active

1-LO_2 — longitudinal optical phonon (middle frequency) i. r. active

1-LO_3 — longitudinal optical phonon (highest frequency) i. r. active

2-TA — transverse acoustical phonon (frequency unknown)

1-LA — longitudinal acoustical phonon (frequency unknown)

Apart from the O_4, TA and LA phonon branches, the frequencies of the other phonon branches at K ≈ 0 at the center of the Brillouin zone have been determined.[6]

The following restrictions were used to assign multiphonon processes to the observed infrared absorption bands and minor oscillators in the measured reflectance spectra.

1. To a first approximation, all crystals studied had a center of symmetry at 85°K. Therefore, only combinations of phonons from different branches were allowed.[1]

2. Due to the large effective charges of the normal modes of the fluoride perovskites[6] and the qualitative illustration of the effect of ionicity on the vibrational spectrum of a diatomic lattice for large effective charge,[5] the phonon frequencies at the edge of the Brillouin zone were as follows:

a.• The frequencies of the transverse optical phonon branches were equal to or

slightly less than their frequencies at $K \approx 0$ at the center of the zone.

b. The frequencies of the longitudinal optical phonon branches were less than their values at $K \approx 0$ at the center of the zone and decreased more than a transverse optical phonon branch when going from the center of the zone to the edge.

c. The frequency of the longitudinal acoustical phonon branch was larger than the transverse acoustical phonon branch.

3. The frequency of the triply degenerate optically inactive phonon branch was unknown. This normal mode was associated, however, with an internal motion of the (BF_3) octahedron and it was assumed to have a frequency similar to the other internal motions of this unit.

There were nine phonon branches at the edge of the Brillouin zone to contribute to multiphonon processes seen as absorption bands in transmission and minor oscillators in reflection. With these restrictions, there were a possible

$$N = \frac{1}{2} \frac{n!}{(n-2)!} = \frac{1}{2} \frac{9!}{(9-2)!} = 36 \tag{1}$$

two-phonon combinations from different branches to explain multiphonon processes.

The procedure followed to assign the observed multiphonon processes to combinations of phonons from the edge of the Brillouin zone for the cubic fluoride perovskites are listed below.

1. Starting with the frequencies of the transverse phonon branches at $K \approx 0$ at the

Table IV-2. Frequencies of phonon branches at the edge of the Brillouin zone (cm^{-1}) (85°K).

	KMgF$_3$	KMnF$_3$	KCoF$_3$	KNiF$_3$	KZnF$_3$	RbMnF$_3$
TO$_1$	162	120	132	149	135	109
TO$_2$	291	191	218	241	195	193
TO$_3$	471	414	445	450	424	386
O$_4$	442	325	365	387	338	311
LO$_1$	181	141	150	150	146	120
LO$_2$	335	241	255	282	216	224
LO$_3$	504	450	*	517	452	434
LA	94	70	90	82	70	90
TA	70	64	81	70	66	80

*Not assigned due to a lack of available multiphonon processes.

center of the zone, these frequencies were adjusted until two-phonon combinations of them agreed with as many observed multiphonon processes as possible consistent with the stated restrictions. The fitted frequencies were then assigned to the frequencies of the transverse optical phonon branches at the edge of the zone.

2. The frequencies of the transverse optical phonon branches at the edge of the zone were subtracted from the remaining observed multiphonon processes. Numbers which repeated in this list of subtractions were chosen as possible frequencies of unknown phonon branches at the edge of the zone. This procedure was followed in choosing the remaining six unknown phonon branch frequencies at the edge of the zone and the frequencies were assigned to phonon branches with the help of the list of restrictions placed on the branches.

Tables IV-2 through IV-5 list the frequencies of the phonon branches at the edge of the Brillouin zone and the assignments determined from the best fit to the observed multiphonon absorptions.

Ferguson, Guggenheim, and Wood[7] have observed vibrational intervals in their study of electronic absorption spectra of the Ni^{++} ion in $KNiF_3$ with thin samples at helium temperatures. In a second paper,[9] they assigned the observed vibrational intervals to combinations of $K \approx 0$ optical (transverse) phonon modes with forbidden electronic transitions. The phonons break down the selection rules for forbidden electric dipole transitions. The first four frequencies observed by them in each of the $^3A_{2g} \rightarrow {}^3T_{1g}^a$, $^3A_{2g} \rightarrow {}^1T_{2g}$, and $^3A_{2g} \rightarrow {}^3T_{1g}^6$ transitions were assigned to combinations of the forbidden electronic transition with one of the four transverse optical phonons at $K \approx 0$ from the center of the Brillouin zone. They used the vibrational intervals as a basis for this assignment. A study of their vibrational intervals and the phonon branch frequencies at $K \approx 0$ at the center of the zone and at the edge of the zone[6] have led to the electronic-vibrational assignments given in Table IV-6.

The frequencies of the phonon branches at the edge and at $K \approx 0$ at the center of the Brillouin zone were for 85°K,[6] while Ferguson, Guggenheim, and Wood[7] made their measurements at 4°K. The value of 469 cm^{-1} assigned to the TO_3 branch at $K \approx 0$ at the center of the zone to fit the vibrational intervals was a reasonable assignment for the frequency of the transverse optical phonon branch of $KNiF_3$ at 4°K since this phonon frequency increased as the crystal was cooled from 300°K to 85°K.[6] The frequencies of the remaining phonons were not changed for the assignments of the vibrational intervals at 4°K and the calculated vibrational intervals for the assignments listed in Table IV-6 were very close to those reported.[7]

The first absorption reported by Ferguson, Guggenheim, and Wood[7] was assigned here to the transverse acoustical mode at the edge of the Brillouin zone, the next four, to the transverse optical phonon modes at $K \approx 0$ at the center of the zone in the order of increasing frequency. This gave a frequency of 387 cm^{-1} for the triply degenerate

Table IV-3. Multiphonon assignments to experimental data (85°K).

KMgF$_3$		KNiF$_3$	
Observed frequency	Assignment and calculated frequency	Observed frequency	Assignment and calculated frequency
403	$LO_2 + TA = 405$	324	$TO_2 + LA = 323$
	†$TO_3 - TA = 401$	354	$LO_2 + TA = 352$
430	$LO_2 + LA = 429$	363	$LO_2 + LA = 364$
	†$LO_3 - TA = 434$	390	$TO_1 + TO_2 = 390$
*499	$TO_1 + LO_2 = 497$	*470	$O_4 + LA = 469$
*516	$LO_1 + LO_2 = 516$	*521	$TO_2 + LO_2 = 523$
*541	$TO_3 + TA = 541$	585	$LO_3 + TA = 587$
615	$TO_3 + TA + TA = 611$	630	$TO_2 + O_4 = 628$
632	$TO_1 + TO_3 = 633$	667	$TO_1 + LO_3 = 666$
667	$TO_1 + LO_3 = 666$	835	$TO_3 + O_4 = 837$
735	$TO_2 + O_4 = 733$	1210	$TO_3 + LO_3 + TO_2 = 1208$
762	$TO_2 + TO_3 = 762$		
794	$TO_2 + LO_3 = 795$		
803	$TO_3 + LO_2 = 806$		
827	$LA + O_4 + TO_2 = 827$		
857	$LA + TO_3 + TO_2 = 856$		
911	$TO_3 + O_4 = 913$		
943	$LO_3 + O_4 = 946$		

*Seen in reflection.
†Difference bands.
All frequencies in (cm^{-1}).

Table IV-4. Multiphonon assignments to experimental data (85°K).

K (50% Mg + 50% Ni) F$_3$

Observed frequency	Assignment and calculated frequency	Pure crystal
407	LO$_2$ + TA = 405	(KMgF$_3$)
*527	TO$_2$ + LO$_2$ = 523	(KNiF$_3$)
	TO$_1$ + O$_4$ = 536	"
	LO$_1$ + O$_4$ = 537	"
	TO$_3$ + LA = 532	"
	TO$_3$ + TA = 520	"
	O$_4$ + LA = 536	(KMgF$_3$)
710	TO$_2$ + TO$_3$ = 691	(KNiF$_3$)
	LO$_2$ + TO$_3$ = 732	"
	TO$_2$ + O$_4$ = 733	(KMgF$_3$)
	LO$_1$ + LO$_3$ = 685	"
875	TO$_1$ + TO$_3$ + LO$_2$ = 882	(KNiF$_3$)
	TA + LO$_3$ + LO$_2$ = 869	"
	LA + LO$_3$ + LO$_2$ = 881	"
	O$_4$ + LO$_2$ + LA = 821	(KMgF$_3$)
	TO$_3$ + LO$_2$ + TA = 876	"
	TO$_2$ + LO$_3$ + LA = 889	"
	LO$_1$ + LO$_1$ + LO$_3$ = 886	"

*Seen in reflection.
All frequencies in (cm^{-1}).

Table IV-5. Multiphonon assignments to experimental data (85°K).

KMnF$_3$

Observed frequency	Assignment and calculated frequency
*261	$TO_1 + LO_1 = 261$
	† $O_4 - TA = 261$
	$TO_2 + LA = 261$
*432	$TO_2 + LO_2 = 432$
*445	$TO_1 + O_4 = 445$
775	$LO_3 + O_4 = 775$

KCoF$_3$

Observed frequency	Assignment and calculated frequency
*231	$LO_1 + TA = 231$
*240	$LO_1 + LA = 240$
350	$TO_1 + TO_2 = 350$
*473	$TO_2 + LO_2 = 473$
810	$TO_3 + O_4 = 810$

KZnF$_3$

Observed frequency	Assignment and calculated frequency
*201	$TA + TO_1 = 201$
*205	$LA + TO_1 = 205$
330	$TO_2 + TO_1 = 330$
351	$LO_2 + TO_1 = 351$
362	$LO_1 + LO_2 = 362$
*473	$TO_1 + O_4 = 473$
640	$TO_3 + LO_2 = 640$
790	$LO_3 + O_4 = 790$

RbMnF$_3$

Observed frequency	Assignment and calculated frequency
*200	$TO_1 + LA = 199$
	$LO_1 + TA = 200$
333	$TO_1 + LO_2 = 333$
*420	$TO_2 + LO_2 = 417$
	$TO_1 + O_4 = 420$
*431	$LO_1 + O_4 = 431$
600	$TO_3 + LO_2 = 610$
745	$LO_3 + O_4 = 745$

*Seen in reflection.

†Difference band.

All frequencies in (cm^{-1}).

Table IV-6.   Assignment of electronic-vibrational spectrum of $KNiF_3$ reported by Ferguson, Guggenheim, and Wood.[7]

Measured vibrational interval (°K)	Assignment	Calculated vibrational interval
	Pure electronic (E)	
	$E + TA(e) = E + 70$	
81	$E + TO_1(c) = E + 150$	80
177	$E + TO_2(c) = E + 248$	178
317	$E + O_4(c) = E + 387$	317
399	$E + TO_3(c) = E + 469$	399
487	$E + TO_2(c) + LO_2(c) = E + 556$	486
546	$E + LO_2(c) + LO_2(c) = E + 616$	546
625	$E + LO_2(c) + O_4(c) = E + 695$	625
709	$E + LO_3(c) + TO_2(c) = E + 782$	712
832	$E + LO_3(e) + O_4(e) = E + 904$	834

(c) Phonon at $K \approx 0$ from center of Brillouin zone.

(e) Phonon from edge of Brillouin zone.

All frequencies in (cm^{-1}).

optically inactive mode, agreeing with the assignment made earlier (see Table IV-2). The remaining vibrational intervals were assigned to multiphonon processes.

E. F. Young, C. H. Perry

References

1. M. Lax and E. Burstein, Phys. Rev. 97, 39 (1955).

2. B. Szigeti, Proc. Roy. Soc. (London) A252, 217 (1959).

3. J. L. Birman, Phys. Rev. 127, 1093 (1962).

4. J. L. Birman, Phys. Rev. 131, 1489 (1963).

5. R. W. Keyes, J. Chem. Phys. 37, 72 (1962).

6. E. F. Young, Ph. D. Thesis, M. I. T., August 22, 1966.

7. J. Ferguson, H. J. Guggenheim, and D. L. Wood, J. Chem. Phys. 40, 822 (1964).

8. G. R. Hunt, C. H. Perry, and J. Ferguson, Phys. Rev. 134, A688 (1964).

9. J. Ferguson and H. J. Guggenheim, J. Chem. Phys. 44, 1095 (1966).

B. FERROELECTRIC "SOFT" MODE IN KTaO$_3$

Far infrared reflectance measurements have been made on single crystals KTaO$_3$ from 40-1000 microns (250-10 cm^{-1}) over the temperature range 12-463°K. The reflectance data were analyzed using a Kramers-Kronig method to obtain total conductivity, σ, from the imaginary part of the dielectric constant. The conductivity exhibits a very strong temperature dependent low-frequency peak which moves from 106 cm^{-1} at 463°K, to 25 cm^{-1} at 12°K and accounts for more than 90 per cent of the static dielectric constant, ϵ_o.

The results indicate a relationship between this low transverse optical mode, ω_{t_1} for $\underline{k} \approx 0$ and the static dielectric constant ϵ_o, such that

$$\omega_{t_1}^2 (T) = A/\epsilon_o(T)$$

where $A = 1.9 \times 10^6$ (cm^{-1})^2. This result can be related directly to the Curie-Weiss law behavior of ferroelectrics and the associated implications have been derived and discussed by several authors.[1-3] Consequently, the experimental verification in the case of KTaO$_3$ helps confirm the theory of ferroelectricity in perovskite type materials which indicate a soft mode lowering its frequency and becoming unstable as the temperature is lowered towards the Curie point. This phenomenon had previously only been observed with any certainty for SrTiO$_3$ using far infrared[4] and neutron techniques.[5]

The infrared reflectance measurements at a 10° angle of incidence were performed on a Michelson spectrophotometer[6] using several beam splitters and a liquid-helium cooled germanium detector. The sample was mounted in a separately evacuable chamber and its temperature was measured with two calibrated thermocouples attached to the crystal. The complete temperature range could be accomplished without disturbing the sample and mount.

The total reflectivity curve from 10-4000 cm^{-1} was measured at 296°K, and also at 126°K but very little temperature dependence of the modes at about 200 cm^{-1}, 550 cm^{-1} and the side band at 760 cm^{-1} was observed. The results at room temperature agreed closely with those of Miller and Spitzer[7] and reference should be made to their work for the complete spectrum. Between 0-10 cm^{-1} the extrapolated curves were in good agreement with the reflectivity values calculated from the static dielectric constant measurements of Rupprecht and Bell[8] and Davis.[9] Beyond 250 cm^{-1}, the reflectance curve measured at room temperature was used with the high temperature and room temperature data, while the 126°K curve was coupled with the measurements at other temperatures in order to evaluate the far infrared complex dielectric constant from the Kramers-Kronig analysis. The marked temperature dependence of the reflectivity curve below 100 cm^{-1} can be clearly seen in Fig. IV-1, together with the weakly

temperature dependent mode at 200 cm^{-1}.

The frequencies of the normal optically active transverse modes were derived from peaks in the conductivity $\sigma(\omega)$, where $\sigma_j = \epsilon''_j \omega/2$ and ϵ''_j is the contribution of the jth resonance to the imaginary part of the dielectric constant.[10]

The frequency dependence of the conductivity with temperature below 170 cm^{-1} is shown in Fig. IV-2 and the positions of the "soft" mode are estimated to be accurate to about ± 5 cm^{-1}. This estimate was obtained by varying the low frequency input data to the K-K analysis over the limit of the error ($\pm 2\%$) in the reflectance measurements.

The temperature dependence of the transverse optical modes are given in Table IV-7.

The static dielectric constant ϵ_o can be written in terms of the Curie law

$$\epsilon_o \propto \frac{1}{T - T_c},$$

and consequently $\omega_{t_1}^2 \propto T - T_c$. Measurements of ϵ_o for KTaO$_3$ at microwave frequencies by Rupprecht and Bell[8] from 80-303°K have indicated a modified Curie law behavior such that $\epsilon_o = \frac{B}{T - T_c} + C$, where $B = 5.989 \times 10^4$ °K, $C = 39.32$ and the Curie temperature $T_c = 2.9$°K. This result is shown by the solid curve in Fig. IV-3 together with the temperature dependence of the low frequency mode. The upper four points fall closely on the
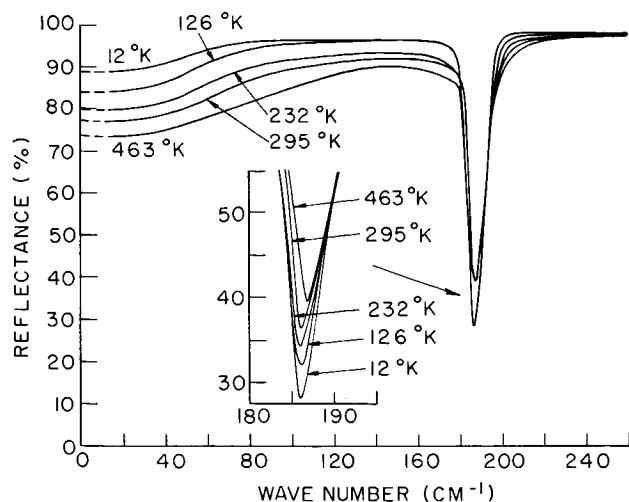

Fig. IV-1. Far infrared reflectivity of KTaO$_3$ as a function of temperature over the frequency range 10-250 cm^{-1}. The error in measurements $\approx \pm 2\%$.
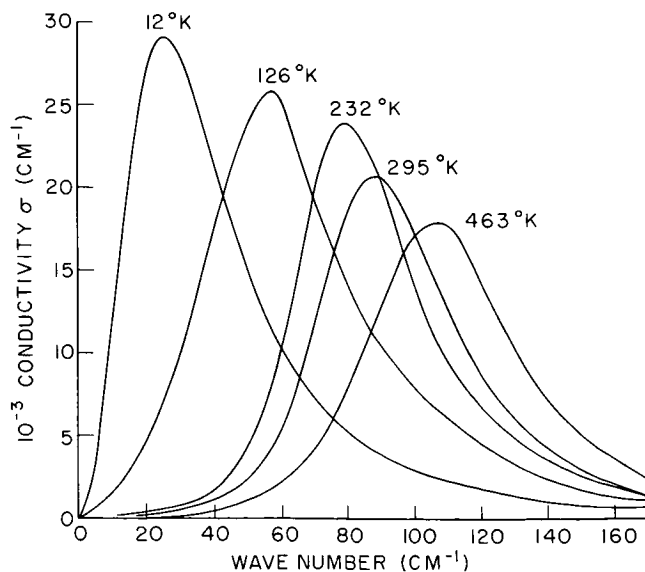

Fig. IV-2. Conductivity calculated from the reflectivity curves shown in Fig. IV-1 for the five different temperatures.
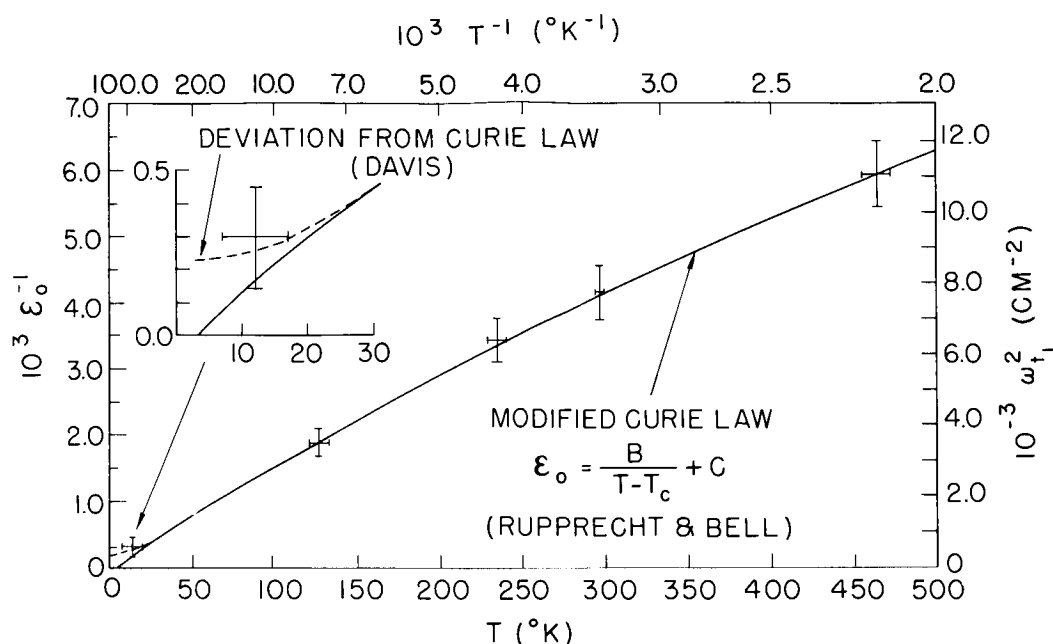
Fig. IV-3. The square of the frequency (in cm^{-1}) of the ferroelectric "soft" mode plotted as a function of temperature. Vertical lines indicate the $\omega_{t_1}^2$ error and horizontal lines show the temperature variation. The solid curve shows the reciprocal of the dielectric constant from the results of Rupprecht and Bell. The extrapolated curves gives a Curie temperature of 2.9°K. The dotted curve shows the low-temperature deviation (below 30°K) from the modified Curie law (R & B) obtained by Davis.

Table IV-7. The transverse optical modes in $KTaO_3$ as a function of temperature.

T°K	ω_{t_1}	ω_{t_2}	ω_{t_3}
12	25	196	—
126	58	198	551
232	79	198	551
295	88*	199*	550*
463	106	199	—

*See Miller and Spitzer[7] for complete spectrum.

modified Curie law described by Rupprecht and Bell but the experimental error in $\omega_{t_1}^2$ barely excludes normal Curie law behavior. Davis[9] has recently measured the dielectric constant below 80°K and has found deviations from the Curie law under 30°K. The one measured frequency in this region (~12°K) also indicates a value corresponding to deviations found by Davis. The extremely low "Curie temperature" makes this crystal ideal for the investigation of the "soft" mode in the paraelectric cubic state. The temperature variation of this vibration is in good agreement with the temperature dependence of the dielectric constant related by the Cochran-Cowley[1,11] theory of ferroeletricity in perovskite crystals.

<div align="right">C. H. Perry, T. F. McNelly</div>

References

1. W. Cochran, Advances in Physics, edited by N. F. Mott (Taylor and Francis, Ltd., London, 1960), Vol. 9, p. 387.

2. P. W. Anderson, Fizika Dielectrikov, edited by G. I. Skanavi (Akad. Nauk S. S. S. R. Fizicheskii Inst. im P. N. Lebedeva, Moscow, 1960).

3. T. Nakamura, The Institute for Solid State Physics, Tokyo, Japan, Ser. A, 186, 1 (1966). J. Phys. Soc. Japan, p. 491 (1966).

4. A. S. Barker, Jr. and M. Tinkham, Phys. Rev. 125, 1527 (1962).

5. R. A. Cowley, Phys. Rev. Letters 9, 159 (1962).

6. C. H. Perry, R. Geick, and E. F. Young, Applied Optics 5, 1171 (1966).

7. R. C. Miller and W. G. Spitzer, Phys. Rev. 129, 94 (1963).

8. G. Rupprecht and R. O. Bell, Phys. Rev. 135, A748 (1964).

9. T. G. Davis, S. M. Thesis, Department of Electrical Engineering, M. I. T. (1966).

10. F. Seitz, Modern Theory of Solids (McGraw-Hill Book Co. , New York, 1940), Chap. XVII, p. 635.

11. W. Cochran and R. A. Cowley, J. Phys. Chem. Solids 23, 447 (1962).

V. PHYSICAL ELECTRONICS AND SURFACE PHYSICS*

Academic and Research Staff

Prof. R. E. Stickney
Dr. J. G. Bergman
Dr. M. L. Shaw

Graduate Students

D. L. Fehrs   R. M. Logan   D. S. Shupe
J. W. Gadzuk   Y. S. Lou   S. Yamamoto

A. SYNTHESIS OF AMMONIA UNDER HIGH-VACUUM CONDITIONS

1. Introduction

The specific objective of this program is to study the role of iron as a catalyst in the synthesis of ammonia by using high vacuum techniques and a mass spectrometer. More generally, we hope to demonstrate the value of high-vacuum surface studies to the field of heterogeneous catalysis.

This first progress report describes the problem, the experimental apparatus, and some preliminary data.

2. Equilibrium

$$\frac{1}{2} N_2 + \frac{3}{2} H_2 \rightleftharpoons NH_3 + heat$$
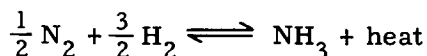
Equilibrium data for this reaction was first obtained by Haber and Van Oordt in 1905[1] and is now established over a wide range of temperatures.[2] The reaction is clearly favored by low temperatures and high pressures.

The experimental equilibrium data can be fitted closely through statistical mechanics by using an energy of formation, ΔE, of -9400 cal/gm-mole; for temperatures in the vicinity of 300 to 600°K, the equilibrium constant varies in proportion to $T^{-3} \exp\left(-\frac{\Delta E}{RT}\right)$, hence increases strongly for a decrease in temperature. However, for an industrially-useful reaction rate, a temperature of at least 625 to 700°K is required. Thus a sharp conflict exists between the demands for a useful rate and for higher production concentrations.

Even if the reaction would proceed at room temperature (this has not been observed), the equilibrium partial pressure of ammonia formed from reactants at pressures directly observable with a spectrometer would be small. For example, at 300°K

with $P_{N_2} = P_{H_2} = 10^{-5}$ Torr, $P_{NH_3} \approx 10^{-10}$ Torr at equilibrium. Yet ammonia was recently synthesized at pressures greater than 10^{-7} Torr under nonequilibrium conditons
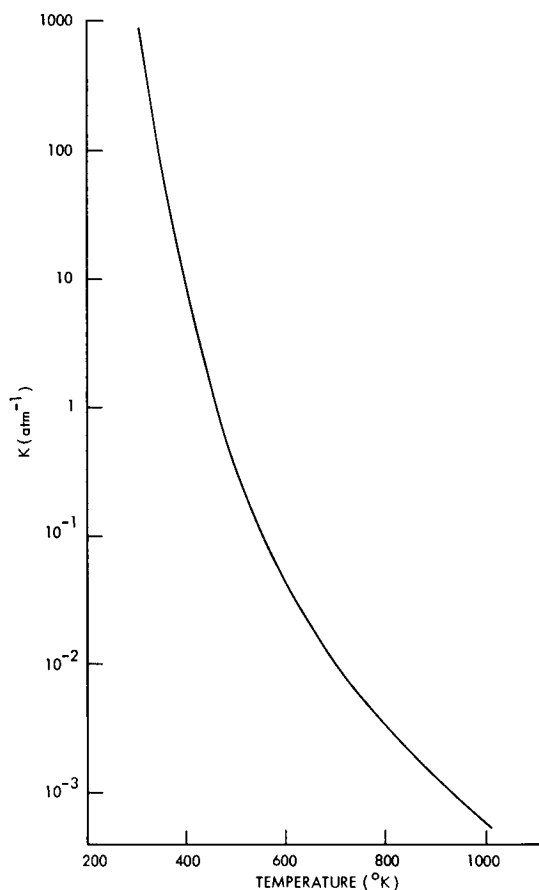
Fig. V-1. Equilibrium constant, K, as a function of temperature for the reaction $\frac{1}{2} N_2 + \frac{3}{2} H_2 \rightleftharpoons NH_3$.

by G. E. Moore who used a molybdenum ribbon as his catalyst.[3] Moore pulsed the ribbon's temperature, thereby probably increasing the surface "pressure" or concentration of one or both of the reactants considerably above equilibrium. He noted that at these low pressures, diffusion effects seem to be of importance comparable to surface effects.

3. Design of Experiment

The heart of the experiment is a thin-walled iron tube commercially produced as a nitrogen diffusion leak for vacuum studies. The heating mandrel and filament normally surrounding the tube have been removed for our study. We reduced the wall thickness of the tube to about 5 mils to allow higher diffusion rates and direct resistance heating.

The mass spectrometer is placed close to the tube so that atoms, molecules, or free radicals leaving the iron surface can travel directly into the analyzer. The temperature
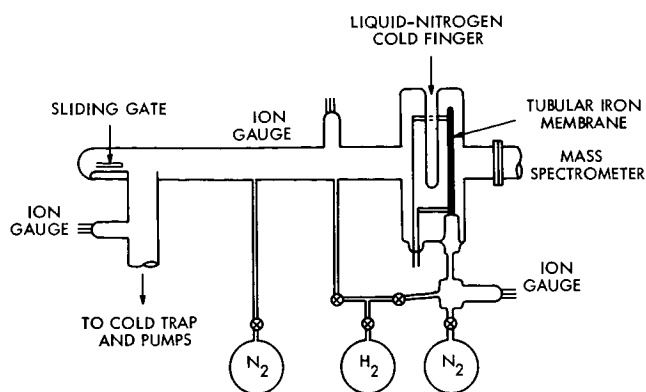
Fig. V-2. Schematic of experimental apparatus for studying NH_3 catalysis.

of the center section of the tube is sensed by a chromel-alumel thermocouple.

Following bakeout and outgassing, pressures of 10^{-9} Torr are obtained with the three-stage mercury diffusion pump. The valves are bakeable Granville-Phillips and the reactant gases are research grade in Pyrex flasks. The mass spectrometer is a General Electric Monopole Partial Pressure Analyzer, Model 163.

With the arrangement shown above, there are four possibilities for studying ammonia synthesis:

1. Both reactants supplied to the outer surface of the tube by chemisorption from the surrounding gas.

2. Both reactants supplied to the outer surface by diffusion from the interior of the iron membrane.

3. and 4. One reactant supplied through chemisorption and other through diffusion.

4. Results

We recently produced ammonia at easily measurable pressures. This was first accomplished by surrounding the iron catalyst, inside and out, with hydrogen and

Fig. V-3.

Spectrum obtained following overnight of pumping with controlled nitrogen and hydrogen leaks. The iron tube was unheated. Total pressure, 3×10^{-7} Torr. Peak 44 is CO_2; peak 28 is N_2 and CO; 19 is fluorine (from the glass); 18 is H_2O; 17 is NH_3 with some OH from H_2O cracking; 16 is mostly NH_2 from NH_3 cracking; 14 is N; 12 is C from CO_2 and CO cracking; and 2 is H_2.

nitrogen at 10^{-4} Torr, then pulsing the membrane while the cold finger was filled with liquid nitrogen. On warming the finger, ammonia desorbed from its surface and was

thereafter present in the system in varying amounts until the next bakeout.

Following bakeout and before breaking the flasks, we unsuccessfully attempted to outgas the iron of CO. Carbon monoxide continues to be produced conspicuously whenever the iron is heated. The tenacity and continued presence of CO with molybdenum and tungsten was recently pointed out as a possible source of error in hydrogen exchange studies not employing a mass spectrometer.[4]

The introduction of hydrogen to the system was accompanied by what appears to be the three lightest paraffins, methane being most prominent. We are now attempting to eliminate these undesirable products before continuing with our investigation of ammonia synthesis.

<div align="right">D. S. Shupe, R. E. Stickney</div>

References

1. F. Haber and G. Van Oordt, "Uber Bildung von Ammoniak aus den Elementem," Z. Anorgan. Chemie 43, 11 (1905).

2. "Thermodynamic Properties of 65 Elements: Their Oxides, Halides, Carbides, and Nitrides," Bulletin 605, Bureau of Mines, C. E. Wicks and F. E. Block, U.S. Government Printing Office, p. 88, 1963.

3. Private communication from George E. Moore, Bell Telephone Laboratories, Murray Hill, New Jersey, 1966.

4. G. E. Moore and F. C. Unterwald, "Interaction of Hydrogen with Tungsten and Molybdenum," J. Chem. Phys. 40, 2625-2652 (1964).

Academic and Research Staff

Prof. K. U. Ingard
Prof. L. W. Dean III

Graduate Students

R. H. Katyl   W. M. Manheimer   J. A. Ross
J. L. Macon   H. M. Schulz III

A. POSSIBILITY OF ION-WAVE AMPLIFICATION IN A PLASMA

The energy-transfer mechanism proposed by Ingard[1] for the amplification of acoustic waves in a weakly ionized gas has been applied to the amplification of ion-acoustic waves in a fully ionized gas. Of importance in the calculation is the assumption, supported by recent experimental observations,[2] that the electron temperature remains constant during the period of an ion-acoustic oscillation. If the ions are treated adiabatically except for the coherent heating by the electrons, and if the various possible ion-acoustic loss terms are ignored, an amplification is discovered which is of order T_i/T_e (i. e., of the order of the ion-electron temperature ratio) times the amplification calculated for acoustic waves in a weakly ionized gas having the same electron density and temperature. This result can be understood from elementary considerations, and a note discussing the calculation in somewhat more detail has been submitted for publication in The Physics of Fluids.

H. M. Schulz III

References

1. U. Ingard, Phys. Rev. 145, 41 (1966).
2. I. Alexeff and W. D. Jones, Phys. Rev. Letters 15, 286 (1965).

B. SOUND AMPLIFICATION IN PLASMAS

Preliminary experimental evidence has been obtained in support of the theory advanced by Ingard[1] that sound waves are amplified as they pass through a plasma. This results from an in-phase heating of the neutrals by the hot electron gas. The experimental configuration is the following. A cold Helium discharge at a pressure of 16 mm Hg was maintained in a long cylindrical tube, 2" in diameter, which contains a speaker at one end and a microphone at the other. The speaker is so constructed that the maximum
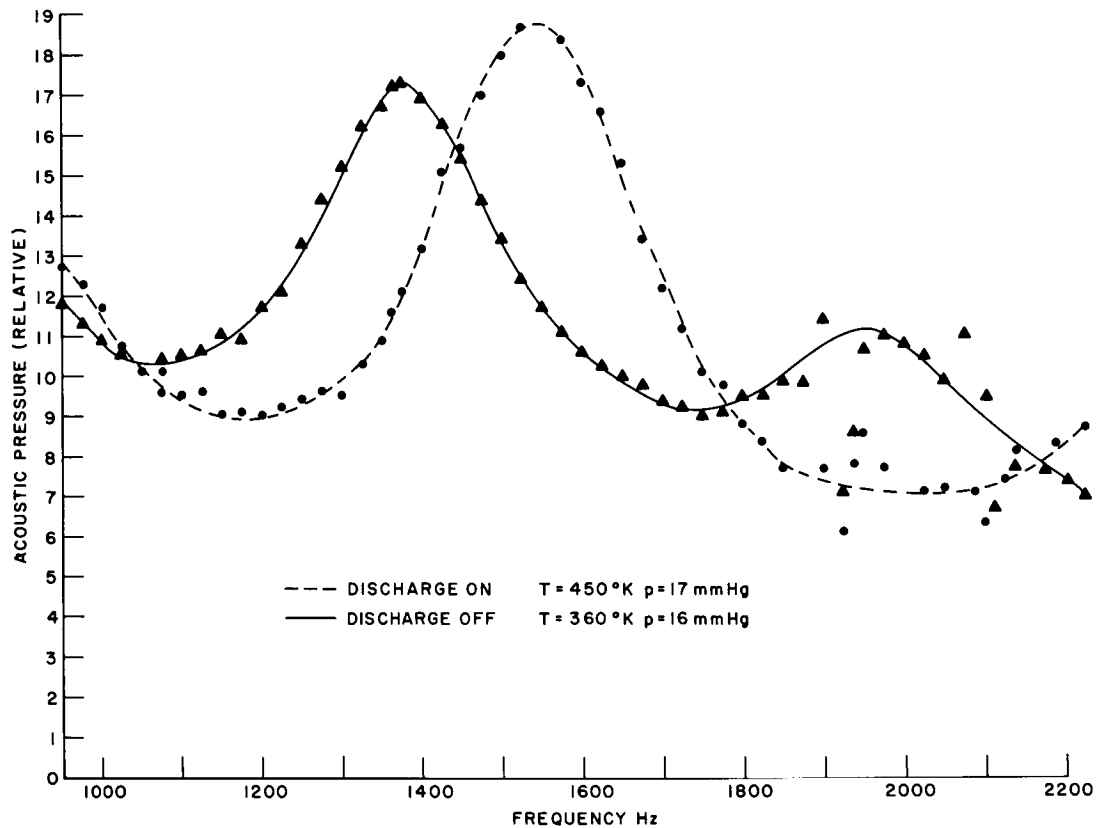
Fig. VI-1. Cavity response vs frequency for a discharge power of 340 watts (E=1600 V, I=210 ma). Pressure is not calibrated.

displacement of the diaphragm can be measured electronically while it vibrates. This known sound pressure source is then used to excite vibrations within the cylindrical cavity, and the resulting pressure at the other end of the tube is measured. The frequencies used are in the 1-2 kc range.

Figure VI-1 gives the results of a typical set of measurements. With the discharge on, the resonance peak for the second normal mode shifted from 1325 Hz to 1540 Hz. This is caused by the increase in neutral temperature when the discharge is turned on. Of significance is the fact that the peak is higher and the preceding trough is lower for the discharge ON curve. This indicates that the net attenuation rate is smaller with the discharge ON than it is with the discharge OFF. This is presumed to result from the electron amplification.

Details of these results are given in the author's thesis.[2]

J. L. Macon

References

1. U. Ingard, Phys. Rev. 145, 41 (1966).
2. J. L. Macon, S.M. Thesis, M.I.T., September 1966.

PLASMA DYNAMICS

Academic and Research Staff

Prof. S. C. Brown Dr. G. Lampis E. M. Mattison
Prof. W. P. Allis J. J. McCarthy W. J. Mulligan
Prof. J. C. Ingraham F. Y-F. Tse

Graduate Students

M. L. Andrews P. W. Jameson D. T. Llewellyn-Jones
D. L. Flannery R. L. Kronquist G. L. Rogoff
E. V. George D. W. Swain

A. ION-WAVE INSTABILITY IN A STEADY-STATE DISCHARGE

Following is a calculation of an ion-wave instability that can occur in a steady-state discharge in which no electric currents exist. The calculation assumes that all other plasma instabilities are absent so that to observe this instability experimentally conditions must be accordingly adjusted.

The plasma is assumed weakly ionized so that electron-ion collisions are negligible. The starting equations are the conservation of particle and momentum equations for electrons and ions and Poisson's equation:

$$\frac{\partial n_e}{\partial t} + \vec{\nabla} \cdot \overrightarrow{n_e v_e} = n_e S(T) \tag{1}$$

$$\frac{\partial}{\partial t} \overrightarrow{n_e v_e} + \frac{\gamma_e kT}{m_e} \vec{\nabla} n_e + \frac{n_e e}{m_e} (-\nabla\phi + \vec{v}_e \times \vec{B}) = -n_e \vec{v}_e \nu_e \tag{2}$$

$$\frac{\partial n_+}{\partial t} + \vec{\nabla} \cdot \overrightarrow{n_+ v_+} = n_e S(T) \tag{3}$$

$$\frac{\partial}{\partial t} \overrightarrow{n_+ v_+} - \frac{n_+ e}{m_+} (-\overrightarrow{\nabla\phi} + \vec{v}_+ \times \vec{B}) = -\overrightarrow{n_+ v_+} \nu_+ \tag{4}$$

$$\nabla^2 \phi = -\frac{e}{\epsilon_o} (n_+ - n_e), \tag{5}$$

where the ion temperature is neglected, and T is the electron temperature. The electrons are assumed to obey an equation of state

$$p_e = p_{eo} \left(\frac{n_e}{n_{eo}}\right)^{\gamma_e}. \tag{6}$$

The electron-atom and ion-atom collision frequencies are ν_e and ν_+, respectively. The particle source term $n_e S(T)$ balances the particle losses resulting from diffusion and recombination in a steady-state discharge. This term is generally a strongly increasing function of electron temperature so that if the electron temperature increases above the value corresponding to the steady state, production exceeds loss, and the ionization density grows. A low-frequency wave which modulates the electron temperature is thus coupled to the production of ionization through this source term.

Assume that the steady-state discharge is governed by ionization production through electron-atom ionizing collisions where the electrons have a Maxwellian velocity distribution. Assume that the ionization loss is due to diffusion to the walls of a cylindrical container of radius R and length L, having a magnetic field parallel to its axis. At steady state the production rate is[1]

$$n_s S(T_s) = n_s A T_s^{3/2} \; e^{-u_i/kT_s} \left(1 + \frac{u_i}{2kT_s} \right), \tag{7}$$

where A is a constant characteristic of the gas, and u_i is its ionization potential. The subscript "s" denotes steady state. The loss term, assuming the density to be distributed in the fundamental diffusion mode, is

$$(\vec{\nabla} \cdot \overrightarrow{n_e v_e})_s = n_s \frac{\mu_+ kT_s}{e} \left(\frac{\left(\frac{2.4}{R} \right)^2}{\left(1 + \mu_+ \mu_e B^2 \right)} + \left(\frac{\pi}{L} \right)^2 \right), \tag{8}$$

where $\mu_+ = \frac{e}{m\nu_+}$ and $\mu_e = \frac{e}{m\nu_e}$ are the ion and electron mobilities, respectively. The equality of (7) and (8) determines T_s in steady-state discharge theory, and corresponds to the zero-order solution of (1) and (4). We assume that a small amplitude wave causes deviations from the steady state: $n_e = n_s + n_{e1}$, $n_+ = n_s + n_{+1}$, $T = T_s + T_1$, and we linearize (1) and (4). We also assume that the tube radius is much greater than the wavelength of the disturbance so that zero-order density gradient effects on the wave propagation are negligible (except as they determine the loss of particles to the walls due to diffusion). This gives

$$\frac{\partial n_{e1}}{\partial t} + n_s \vec{\nabla} \cdot \vec{v}_e = \frac{\partial n_{+1}}{\partial t} + n_s \vec{\nabla} \cdot \vec{v}_+ =$$

$$n_s \left(-\frac{\mu_+ kT}{e} \left[\frac{\left(\frac{2.4}{R} \right)^2}{1 + \mu_+ \mu_e B^2} + \left(\frac{\pi}{L} \right)^2 \right] + A T^{3/2} \; e^{-u_i/kT} \left(1 + \frac{u_i}{2kT} \right) \right), \tag{9}$$

where $\vec{v}_e$ and $\vec{v}_+$ are due only to the wave and not to the density gradients. Using Eqs. 7 and 8 and the fact that they are equal for $T = T_s$, we linearize the right-hand side of Eq. 9 to obtain

$$\frac{\partial n_{e1}}{\partial t} + n_s \vec{\nabla} \cdot \vec{v}_e = \frac{\partial n_{+1}}{\partial t} + n_s \vec{\nabla} \cdot \vec{v}_+ =$$

$$n_s \frac{\mu_+ kT_s}{e} \left(\frac{\left(\frac{2.4}{R}\right)^2}{\left(1 + \mu_+ \mu_e B^2\right)} + \left(\frac{\pi}{L}\right)^2 \right) \left(-1 + \frac{\frac{3}{2} + \frac{5}{4}\frac{u_i}{kT_s} + \frac{u_i^2}{2(kT_s)^2}}{\left(1 + \frac{u_i}{2kT_s}\right)} \right) \frac{T_1}{T_s}.$$

Using $p_e = n_e kT$ and Eq. 6, we obtain $T_1/T_s \cong (\gamma_e - 1) n_{e1}/n_s$, and assuming $u_i/kT_s \gg 1$, Eq. 10 becomes

$$\frac{\partial n_{e1}}{\partial t} + n_s \vec{\nabla} \cdot \vec{v}_e = \frac{\partial n_{+1}}{\partial t} + n_s \vec{\nabla} \cdot \vec{v}_+ = \beta n_{e1}, \tag{11}$$

where

$$\beta \equiv (\gamma_e - 1) \frac{\mu_+ u_i}{e} \left(\frac{\left(\frac{2.4}{R}\right)^2}{1 + \mu_+ \mu_e B^2} + \left(\frac{\pi}{L}\right)^2 \right). \tag{12}$$

The coefficient β will lead to growth of low-frequency ion waves under conditions determined below. In solving the ion-wave problem we assume that

$$n_{e1} = \frac{\epsilon_o \phi}{e \lambda_D^2}, \tag{13}$$

where

$$\lambda_D^2 = \frac{\gamma_e kT_s \epsilon_o}{n_s e^2}. \tag{14}$$

Equation 13 is valid, provided the electrons can move freely along the magnetic field lines so as to neutralize the ion space charge. With this assumption and assuming $\vec{B} = \vec{B}_o$ we are left with only (3), (4), and (5) to solve with n_{+1}, $\vec{v}_+$, and ϕ the unknowns. Assuming that these quantities vary as $\exp\left[i(\omega t - \vec{k}_\perp \cdot \vec{r} - kz)\right]$, where the z-axis is parallel to and $\vec{k}_\perp$ perpendicular to $\vec{B}_o$ we obtain the dispersion relation

$$k_\perp^2\lambda_D^2\left(1 + \frac{\omega_{p+}^2}{i\omega(i\omega+\nu_+)\left(1 + \frac{\omega_{B+}^2}{(i\omega+\nu_+)^2}\right)}\right) + k^2\lambda_D^2\left(1 + \frac{\omega_{p+}^2}{i\omega(i\omega+\nu_+)}\right), + \left(1 - \frac{\beta}{i\omega}\right) = 0 \qquad (15)$$

where $\omega_{p+}^2 = \frac{n_s e^2}{m_+ \epsilon_o}$ is the square of the ion plasma frequency and $\omega_{B+} = \frac{eB_o}{m_+}$ is the ion cyclotron frequency. Assuming $\omega = \omega_r + i\omega_i$ and that k and $k_\perp$ are real allows us to determine the stability of the system, negative ω_i giving rise to instability. The equation in its general form is very difficult to solve and only one example will be treated at this time — that of ion-acoustic wave propagation parallel to $\vec{B}_o$ ($k_\perp = 0$). We obtain from (15)

$$\omega_i = \nu_+ - X + \sqrt{X^2 - \nu_+ X - \omega_r^2}, \qquad (16)$$

where

$$X \equiv \frac{k^2\lambda_D^2\omega_{p+}^2}{\beta}.$$

For $\omega_i = 0$ we obtain from Eq. 15

$$\omega_r^2 = \frac{k^2\lambda_D^2\omega_{p+}^2}{1 + k^2\lambda_D^2} - \nu_+^2. \qquad (17)$$

Substituting (17) in (16), we find that for $\omega_i \leq 0$ we must have

$$\beta \geq \nu_+\left(1 + k^2\lambda_D^2\right) \approx \nu_+ \qquad (18)$$

By using Eq. 18 combined with Eq. 12, it is then possible for a given experiment to calculate at what pressure the first ion acoustic wave instabilities should appear, and from Eq. 17 it is possible to calculate their frequencies.

The damping for waves with $k_\perp \neq 0$ will be greater because of the ion cyclotron absorption, so that as pressure is decreased in a given experiment, the waves with $k_\perp = 0$ will become unstable first.

J. C. Ingraham

References

1. S. C. Brown, Basic Data of Plasma Physics (John Wiley and Sons Ltd., New York, 1959), p. 289.

B. LASER BREAKDOWN EXPERIMENT

The laser breakdown experiment is proceeding (see Quarterly Progress Report No. 81, page 58). The pulse from the ruby laser has been stabilized to the power of 50 MW (2 J in 40 nsec). The gas is helium at 10. 25 kg/cm^2 pressure.

The first part of the experiment was devoted to the determination of the transmission rate of the gas laser power (HeNe) through the plasma, with the aim of gaining knowledge of the absorption coefficient of the light at 6328 Å, and measuring the dimensions of the plasma during its evolution in time.

In measuring the transmission of the laser light, it is necessary to take care that the plasma acts as a negative lens on the gas-laser beam, spreading it out. To collect the beam while it is being spread out, and to measure the solid angle of the spreading in time, the geometry of the optics of the early measurements was modified (Fig. VII-1).

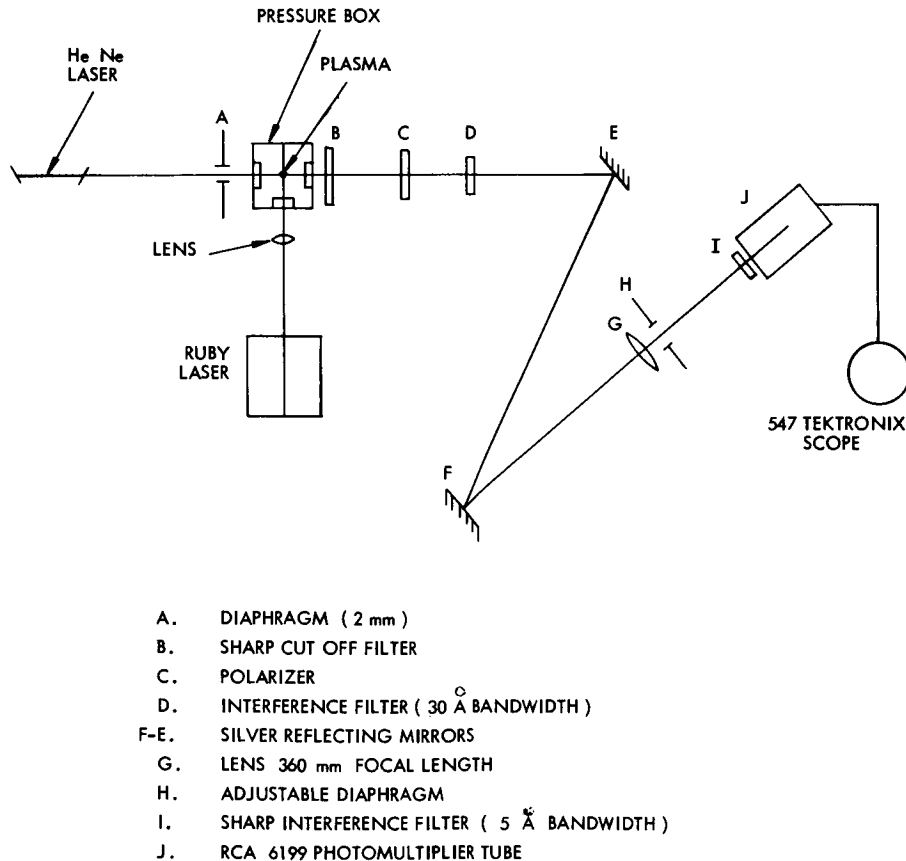|	
A.	DIAPHRAGM (2 mm)
B.	SHARP CUT OFF FILTER
C.	POLARIZER
D.	INTERFERENCE FILTER (30 Å BANDWIDTH)
F-E.	SILVER REFLECTING MIRRORS
G.	LENS 360 mm FOCAL LENGTH
H.	ADJUSTABLE DIAPHRAGM
I.	SHARP INTERFERENCE FILTER (5 Å BANDWIDTH)
J.	RCA 6199 PHOTOMULTIPLIER TUBE

Fig. VII-1. Experimental arrangement for measuring density and absorption.

The diaphragm, A, brings the gas laser beam to a known diameter (2 mm). The sharp cutoff filter, B, eliminates the unwanted light from the plasma with frequency

over 5800 Å, while the rotatable polarizer, C, is set for maximum transmission for the already polarized gas-laser light, thereby favoring by a factor of two the competition between the gas laser light and the plasma emission (which is not polarized), at the same frequency within the band of the two interference filters, D and I, in cascade, both being tilted to peak transmission at 6328 Å.

The diaphragm, H, beyond the lens, G, (focal length 360 mm) can be adjusted in seven positions (f/5.6, f/8, f/11, f/16, f/22, f/32, 3 mm).

The photomultiplier tube is set in such a position that the image of the plasma through the lens, G, is focused on the photosensitive cathode.

The silvered mirrors, E and F, are used to extend the optical path up to the desired length. The optical path was optimized at a length of 183 cm between plasma and lens. Shorter paths favor the plasma emission toward the gas laser; longer paths cause the largest aperature of the diaphragm, I, to catch only a part of the light of the gas laser, which is spread out as a cone by the plasma lens and thus gives a false value of the transmission of the gas laser through the plasma.

With this arrangement, several sets of emission curves (plasma alone) and combination curves (gas laser and plasma) were taken, with different apertures of the diaphragm, H, used. Figures VII-2 and VII-3 show four different diaphragm situations; the geometry of the detection always allowed signals limited enough to prevent the photomultiplier tube from working in the nonlinear region.

From every combination curve, the corresponding emission curve was subtracted; in this way, a set of curves of transmission versus time of the gas laser was obtained for different solid angles of detection.

Figure VII-4 shows the set of absorption curves, for different times, between 175 and 630 nsec after breakdown. The quantity plotted on the ordinate is the transmission per unit solid angle; the quantity on the abscissa is the solid angle itself. (The solid angle is the one through which the adjustable diaphragm sees the plasma.)

It is clearly seen by the set of curves that at every time there is a solid angle where the curve becomes a straight line set at 45°; this means that all diaphragms on the 45° line get the whole laser beam (at the same time) and all of them measure the true transmission, while the smaller diaphragms out of the 45° line get only a part of the laser beam.

For longer times the plasma density slows down and the gas-laser beam approaches its natural divergence, and smaller diaphragms can catch it; this is shown by the movement of the knee on the transmission curves. Plotting the solid angle of the knee against time (Fig. VII-5), we find an exponential decay.

From the theory of a cylindrical plasma lens, the index of refraction of the

ZERO LEVEL

t = 0

f / 5.6

(a)

ZERO LEVEL

t = 0

f / 8

(b)

ZERO LEVEL

t = 0

f / 11

(c)

ZERO LEVEL

t = 0
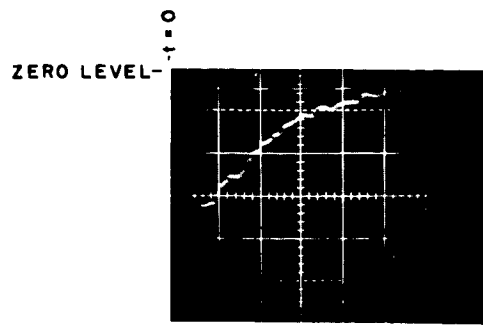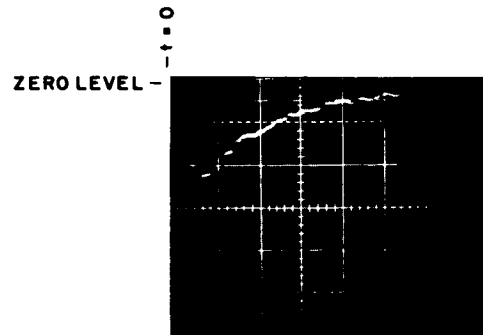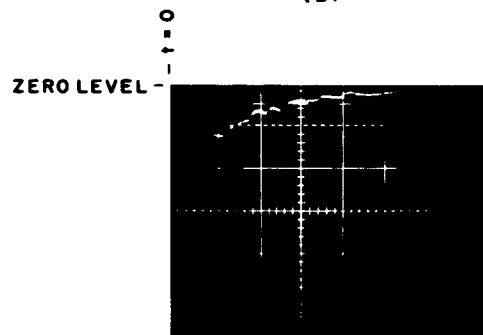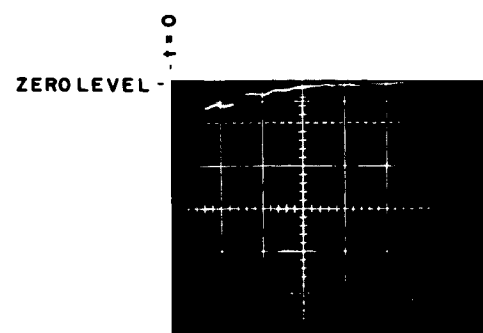
f / 16

(d)

Fig. VII-2.   Plasma emission. Vertical scale: 5 mV/div.
Horizontal scale: 100 nsec/div.

ZERO LEVEL

GAS LASER
STEADY

f/5.6

(a)

ZERO LEVEL

GAS LASER
STEADY

f/8

(b)

ZERO LEVEL

GAS LASER
STEADY

f/11

(c)

ZERO LEVEL

GAS LASER
STEADY

f/16

(d)
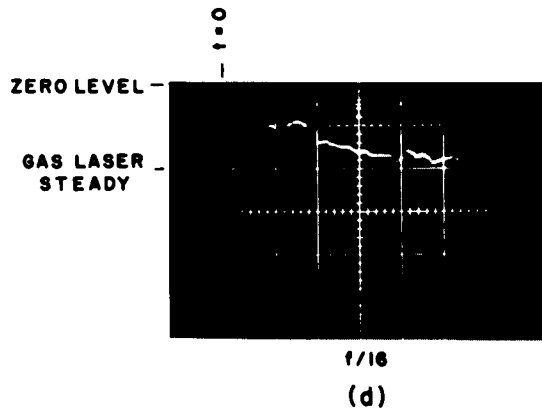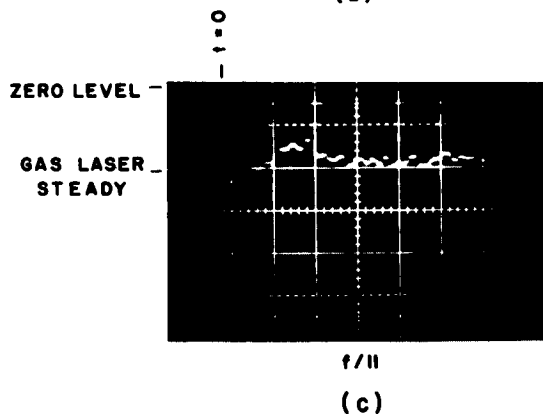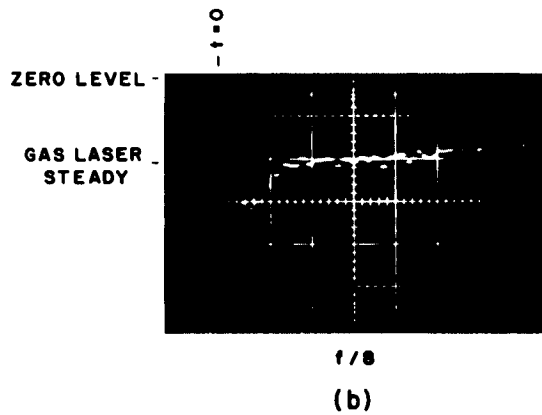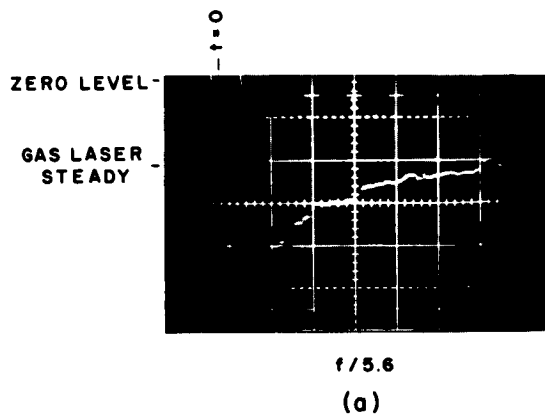
Fig. VII-3.

Combination curves (gas laser plus plasma emission). Vertical scale: 5 mV/div. Horizontal scale: 100 nsec/div.
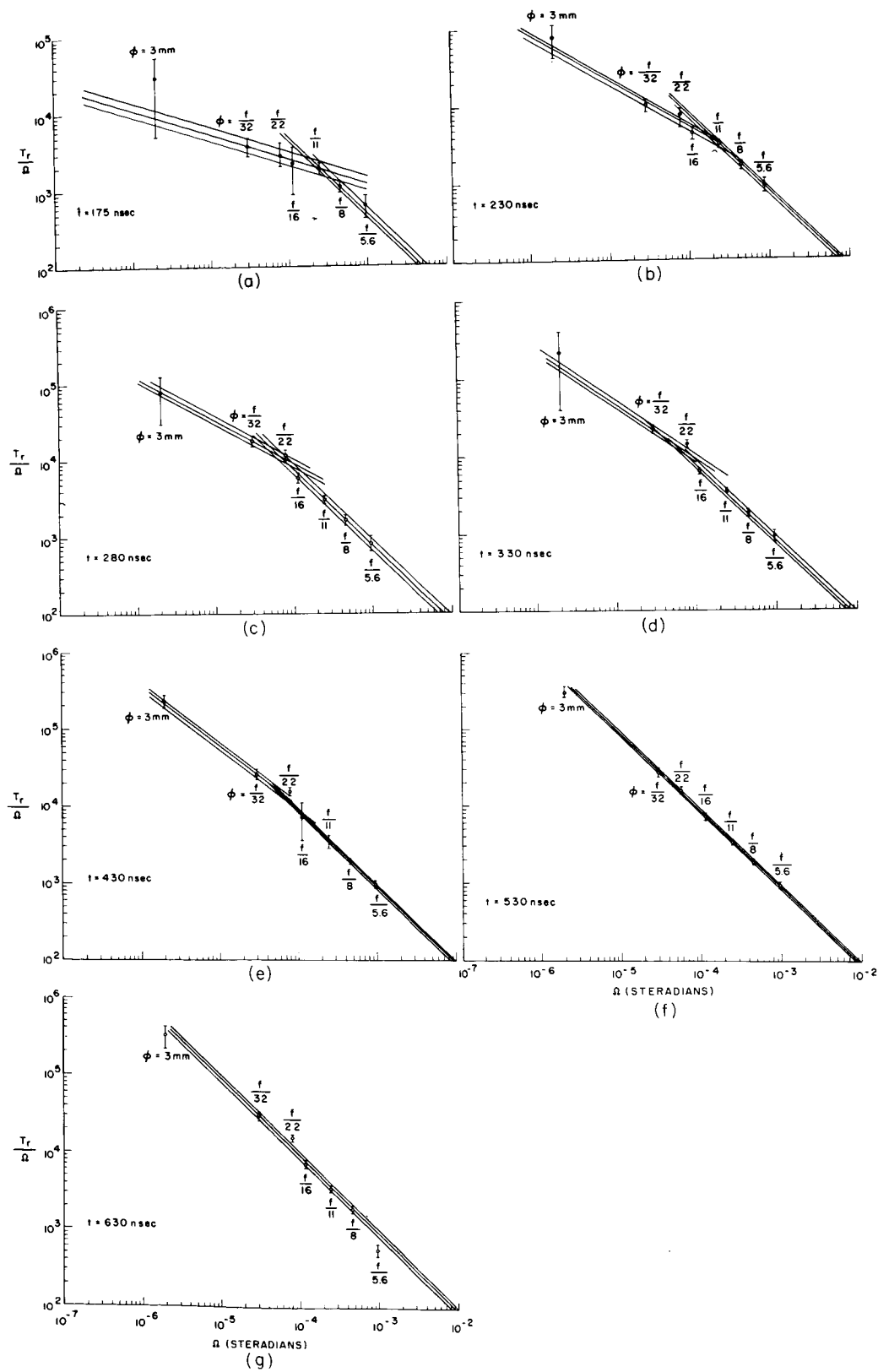
Fig. VII-4. Evolution of the gas-laser spreading and the transmission.

plasma is then immediately known in time (under the assumption that the plasma is uniform.)

$$\frac{1}{r} = 1 + \frac{\theta_e - \theta_i}{2\left(\frac{P}{D} + \theta_i\right)},$$

where θ_e is the half-divergence of the gas laser as it comes out of the plasma (radiants), r is the index of refraction, P is the diameter of the incoming gas laser beam (2 mm), D is the diameter of the plasma (varying in time), and θ_i is the half-divergence (radiants) of the gas laser before entering the plasma.

From the index of refraction, the density follows immediately:

$$r = \left[1 - \left(\frac{\omega_p}{\omega}\right)^2\right]^{1/2} \quad ; \quad \omega_p = 56.45 \cdot 10^3 \cdot N^{1/2},$$

where ω is the frequency of the gas laser light, $2.975 \cdot 10^{15} \frac{rad}{sec}$, ω_p is the plasma frequency, and N is the density $\frac{electrons}{cm^3}$.

Fig. VII-5. Gas-laser spreading decay.

Figure VII-6 shows that the behavior of the density vs time from 100-500 nsec after breakdown is an exponential decay. The density is known within 30 per cent.

Keeping in mind that at a pressure of 10.25 kg/cm^2 the atom density is $2.5 \cdot 10^{20}$ in $\frac{atoms}{cm^3}$, we find at 100 nsec an average ionization degree of 23 per cent.

To find the density for earlier times, it is necessary to use a shorter optical path and thus catch the beam when it is more spread out. This could require a more powerful gas laser, having a power comparable with the emission from the plasma.

A correction of the gas-laser absorption curves was necessary because the gas-laser beams coming out of the lens meet the sharp filter at slightly different angles and result in different transmissions; by shining the gas laser alone, and moving it through known patterns on the surface of the lens at different known angles of incidence, a map of the transmission on all of the surface of the filter was made, thereby yielding a correction factor for every measurement taken with a different diaphragm size.

The error bars in the knee curves were obtained by taking different sets of data and

Fig. VII-6. Plasma-density decay.
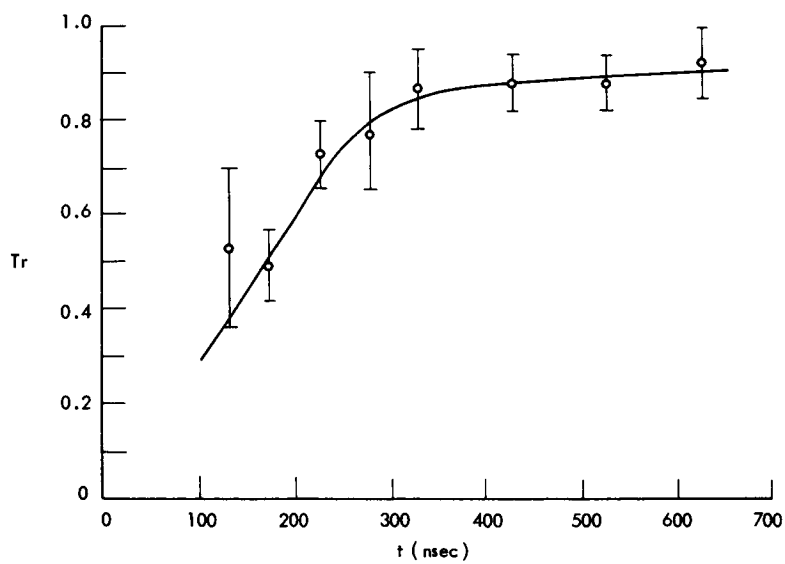
Fig. VII-7. Gas-laser transmission.
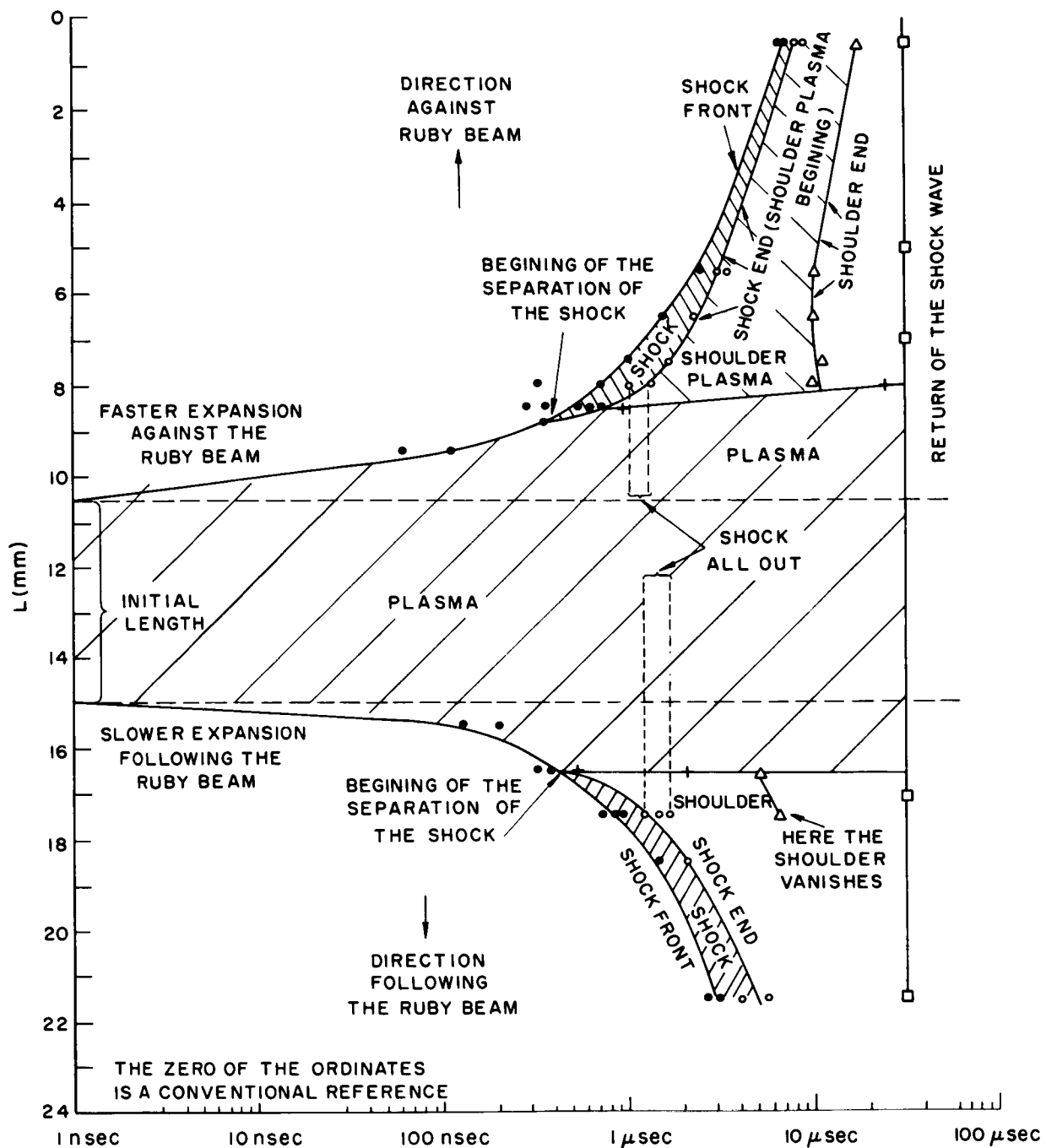
Fig. VII-8. Axial expansion of the plasma probing beam. Beam diameter: 1 mm. Ruby laser power: 50 MW.

keeping account of the noise of the gas laser. From the 45° points of the "knee-curves," it is possible to read the behavior of the transmission of the gas laser through the plasma as a function of time (Fig. VII-7); this quantity, together with the radius of the plasma, gives the absorption coefficient, which is necessary to get the temperature from the emission curves calibrated with the gas laser.[1]

To find the behavior of the length of the plasmas in time, small diaphragms (1-mm diameter) were inserted before and after the windows of the box where the plasma is produced; this gives the gas laser a known section and cuts down the emission from the plasma, so that only the absorption of the gas laser can be detected. The plasma was moved by known amounts along its axis with respect to the gas laser by moving the focussing lens along the ruby beam.

The delay in the beginning of the absorption related to the position of the plasma gives the information on the axial expansion.

Figure VII-8 shows the behavior of the length of the plasma up to 32 μsec. At ~500 nsec the plasma starts to throw out a shock wave from the two heads which produces a plasma itself after its own shoulder. Figure VII-9 shows the instant when the shock separates from the plasma (side of the plasma toward the lens), 1.1-1.3 μsec. Figure VII-10 shows the shock already separated from the initial blob and dragging the shoulder plasma.

The initial length of the plasma is 4.5 mm; it grows to 7.9 mm in 500 nsec. While emitting the shock, the plasma cools strongly and the expansion rate slows down; at 1 μsec the length is 8 mm.

At 32 μsec after breakdown the shock waves emitted from the two heads of the cylinder of plasma are reflected back on the plasma by the walls of the box, and the plasma becomes enormously turbulent and inhomogeneous.

Fig. VII-9. Shock wave from the head of the plasma toward the lens. (Direction against the ruby beam.) Horizontal scale: 500 nsec/div. Vertical scale: 5 mV/div.

It can be seen from Fig. VII-8 that the plasma explodes mainly toward the focussing lens (direction against the ruby beam); the shock-wave velocity and the axial-expansion velocity of the plasma are faster in the direction against the lens, as predicted by the theory.[2] Also, the time when the shock is completely emitted is different from one head to the other head of the plasma, as seen from Fig. VII-8. Further work is in progress on this point.

The fact that the shock surrounding the plasma drags on ionized layer (shoulder plasma) warns us that in any temperature measurement made when this plasma is present (from 500 nsec after breakdown) the emission of this shoulder plasma itself should be taken into account. Figure VII-8 shows, too, the border of existence of this "shoulder" on both sides of the initial cylinder of plasma.

To measure the radial expansion of the plasma, a new geometry was set up (Fig. VII-11). By tilting the quartz slab before the box, it was possible to offset up and down the gas-laser beam with a precision up to 1/10 mm. The gas laser was then reset for every position with a second slab after the box. On the two sides of the box two vertical slits (1 mm wide, 2 cm long) were set up to let the gas-laser beam through and cut down the emission, and a small diaphragm was put before the detector (2-mm diameter). Again, the delay in the beginning of the absorption related to the position of the beam gives the information on the evolution in time of the radius of the plasma. The initial position is found when the plasma explodes at the surface of the gas-laser beam; by offsetting the beam of the first 4/10 mm, the first delay of 30 nsec is seen. Figure VII-12 shows the behavior of the radius growth from the initial value of 0.15 mm (focal spot of the lens) as a function of time.
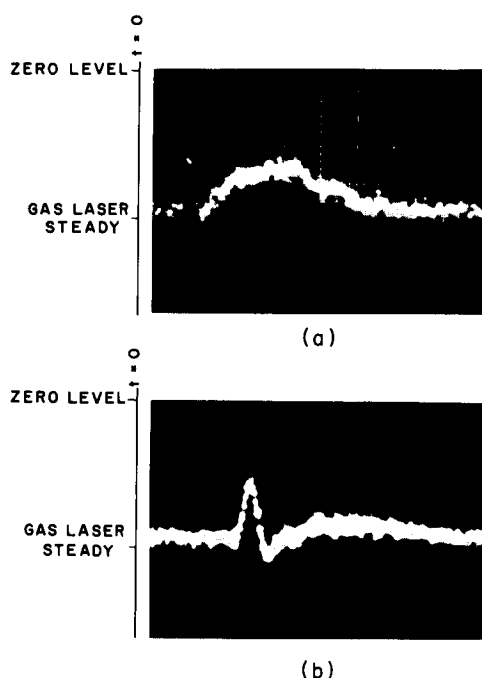
Fig. VII-10. Shock wave dragging its own shoulder plasma.

The plasma throws out a radial shock wave whose position is shown, together with the position of its end, coincident with the beginning of the shoulder plasma. The separation of the shock from the plasma starts around 300 nsec after breakdown, and the shock is completely separated 1.2-1.5 μsec after breakdown.

The knowledge of the radius of the plasma is indispensable, in getting the index of refraction from the measurements of its lens effect on the gas-laser beam, and the absorption coefficient from the curve of transmission. When using the gas laser, both the length and the radius of the plasma are indispensable to compute the temperature from the calibrated emission curves and the absorption coefficient.

PRESSURE BOX

90 cm

He Ne LASER

A
B
C
C
D
A
E
G
H
F

PLASMA   LENS

RUBY LASER

547 TEKTRONIX SCOPE
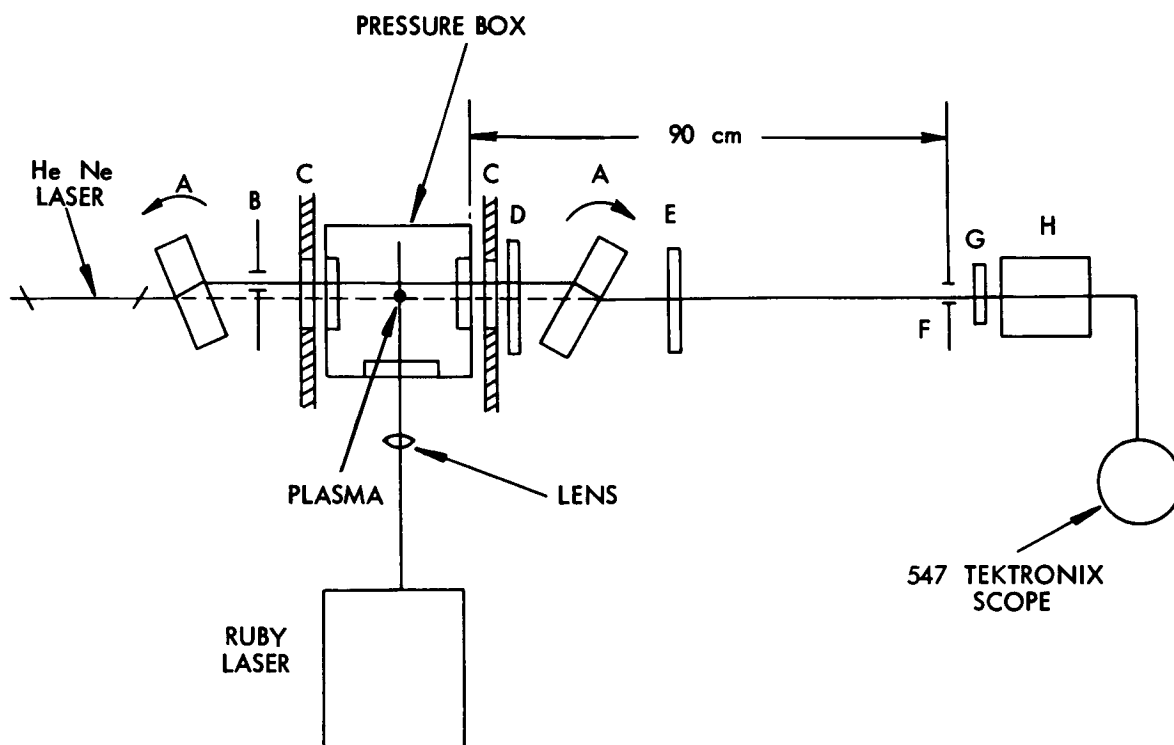
A.   QUARTZ TILTABLE SLABS
B.   VERTICALLY MOVABLE DIAPHRAGM
C.   VERTICAL SLITS (1 mm. WIDE, 2 cm. LONG)
D.   SHARP CUT OFF FILTER
E.   POLARIZER
F.   DIAPHRAGM (2 mm)
G.   SHARP INTERFERENCE FILTER
H.   RCA 6199 PHOTOMULTIPLIER TUBE

Fig. VII-11. Experimental arrangement for measuring the radial expansion of the plasma.
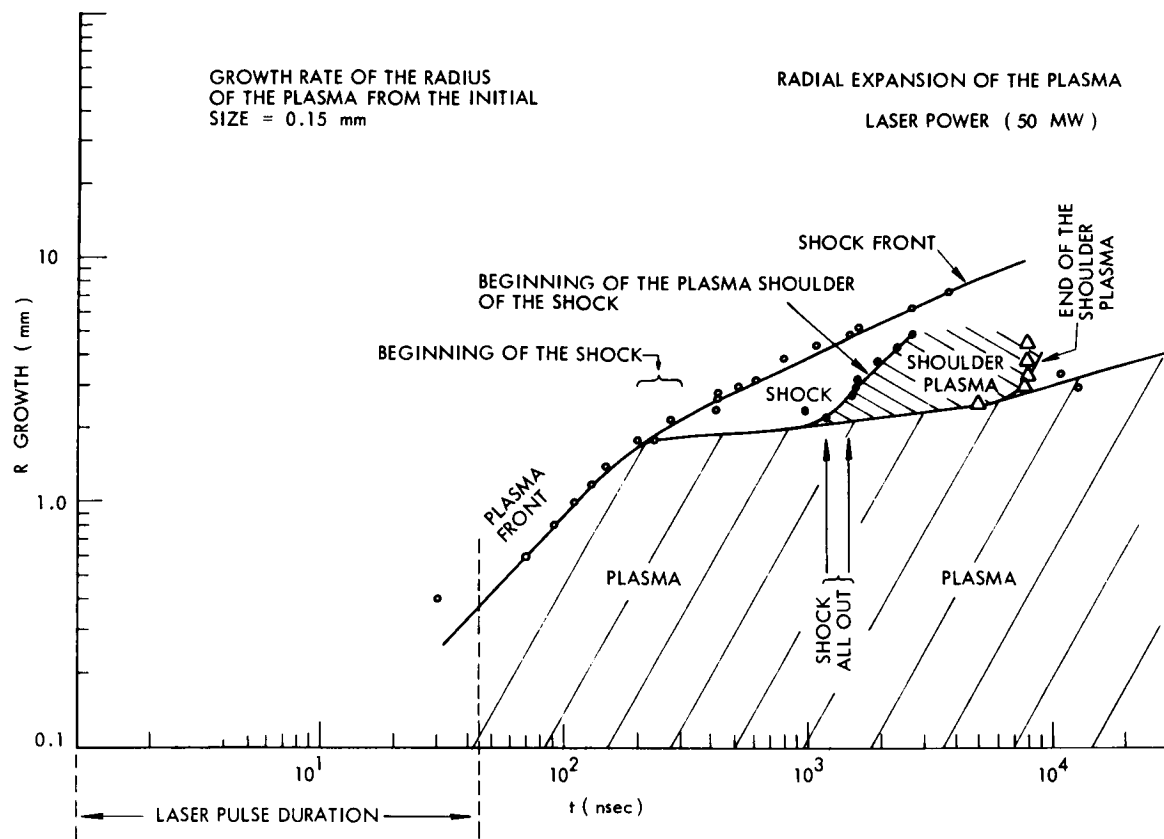
Fig. VII-12. Evolution of the radial expansion of the plasma.

The calibration of the emission curves taken with different diaphragm apertures, for optimizing with respect to angle effects on the sharp filter, has been carried out and computation of the temperature of the plasma is in progress.

G. Lampis

References

1. G. Lampis, "Laser Breakdown," Quarterly Progress Report No. 81, Research Laboratory of Electronics, M.I.T., April 15, 1966, pp. 58-61.

2. Yv. P. Raizer, Soviet Phys. — JETP, Vol. 21, No. 5, November 1965.

3. Yv. P. Raizer, Soviet Phys. — Usp., Vol. 22, No. 5, March-April 1966.

C. MICROWAVE SCATTERING FROM STANDING PLASMA WAVES

Observations of microwave scattering from density fluctuations in an electron-beam produced plasma[1] have continued. The nature of the fluctuations responsible for the scattering now seems quite clear. All of the experimental evidence indicates that the scattering is from plasma waves excited by the electron beam which reflect from the ends of the plasma column and form standing waves along the axis of the tube.

The experimental geometry and procedure are similar to those previously reported[1]; this report also contains references to the theory of density fluctuations and the scattering of electromagnetic waves from these fluctuations.

The present experiments were carried out in a plasma produced by firing an electron beam into un-ionized mercury vapor. The plasma tube was sealed off, and the vapor pressure of the mercury was determined by the temperature of the reservoir of liquid mercury. In all cases this temperature was held at $0°C$, with the result that there was a vapor pressure of approximately 2×10^{-4} mm Hg. The electron gun was a focussed Pierce-type gun. The axis of the tube was aligned along a uniform magnetic field, but the gun was shielded from the field, so that the transverse energy of electrons emerging from the gun was increased as they entered the magnetic field. Since the beam electrons have some transverse energy, the axial velocity is expected to be close to, but somewhat
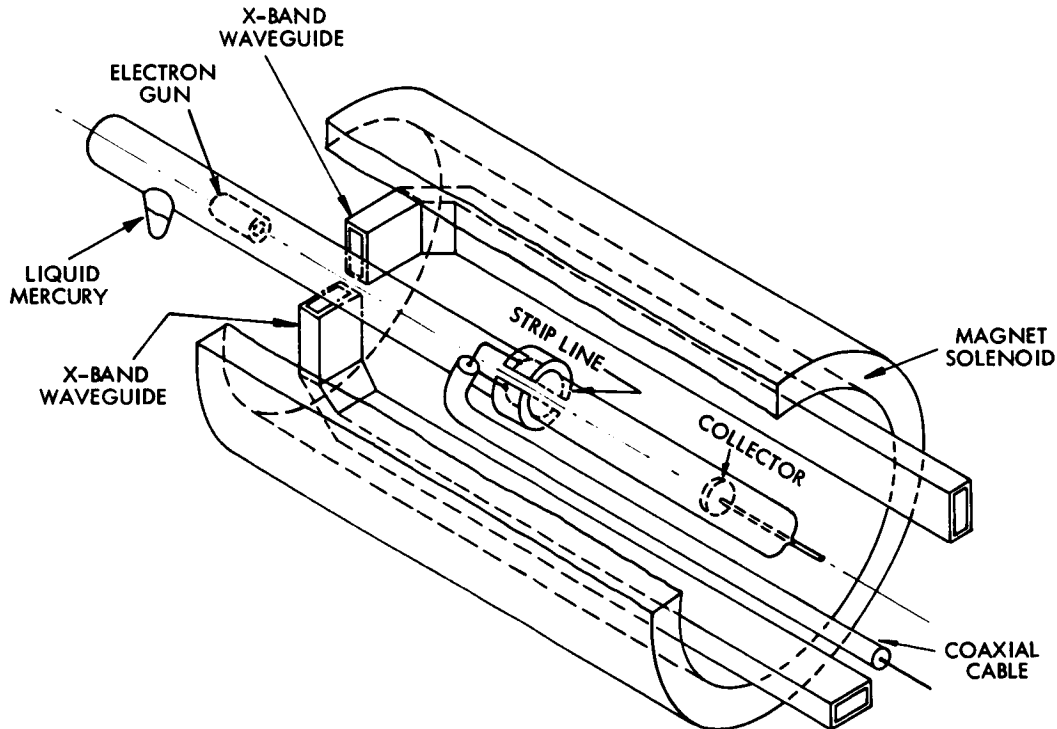
Fig. VII-13. Experimental geometry.

less than, the total beam velocity determined by the voltage applied to the electron gun.

The experimental geometry is shown in Fig. VII-13. Two open-ended pieces of X-band waveguide, which serve as microwave horns, enter the magnet solenoid and bend 90° to face the plasma tube. The horns are oriented so that the electric field for the lowest waveguide mode is parallel to the magnetic field. The horns are mounted on a movable structure so that their position along the axis of the tube can be varied. Displaced 11 cm along the axis from the horns is a strip-line antenna that couples capacitively to the plasma and picks up the plasma oscillations directly. The strip line is also mounted on the movable structure and moves along the axis with the horns.

The microwave circuit used in the scattering experiments is shown in Fig. VII-14.
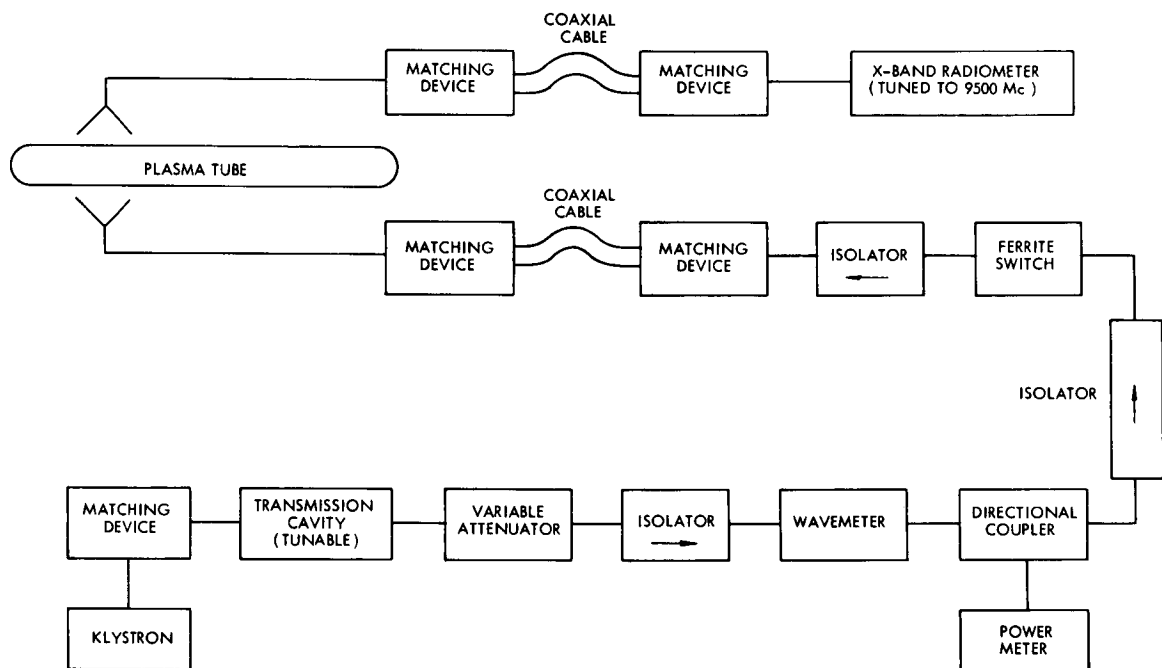
Fig. VII-14. Waveguide circuit.

Microwave radiation at frequency f_{inc} is incident on the plasma from one of the microwave horns. Scattered radiation is picked up by the other horn and fed to an X-band radiometer tuned to 9500 Mc. The amplitude of the scattered radiation at frequency $f_{scat} = 9500$ Mc is approximately proportional to the potential energy in the component of density fluctuations at the difference frequency $f_{scat} - f_{inc}$.

The scattered power was found to be roughly proportional to the incident power, which verified the fact that the effect is scattering of the incident radiation by the plasma, rather than direct emission at the radiometer frequency from the plasma. The scattered power was also found to vary greatly with the frequency of the incident radiation.

The electron beam is expected to interact with the plasma by the familiar two-stream instability mechanism to set up plasma waves along the axis of the tube, which could be reflected from the collector end of the tube and form standing waves. An attempt was therefore made to determine if standing waves existed along the axis of the tube. With the beam voltage and incident microwave power and frequency kept constant, the horns were moved along the axis of the tube. The support structure for the horns was moved by a screw and gear system, which also turned a small 10-turn potentiometer. This potentiometer was used to drive the x axis of an x-y recorder, and the output of the X-band radiometer was applied to the y axis of the recorder. The scattered power was then plotted against the axial position of the microwave horns. The result is shown in Fig. VII-15. The maxima and minima of the scattered power correspond to the maxima and minima of the density fluctuations, and these in turn represent the antinodes and nodes, respectively, of a standing plasma wave. The wavelength λ can be determined from the relation, $d = \lambda/2$, where d is the distance between nodes. The frequency of the standing wave is the difference between the frequencies of the scattered and incident microwaves. The phase velocity given by $v_p = \lambda f$ can then be calculated. We find that $\lambda = 3.6$ cm, $f = 232$ Mc, $v_p = 8.4 \times 10^8$ cm/sec. This is just 5 per cent less than the total beam velocity, which is 8.8×10^8 cm/sec. It seems quite clear that the scattering is from standing plasma waves excited by the electron beam.

$V_{beam} = 220$ volts

$P_{inc} \approx 300$ mW

$f_{inc} = 9268$ Mc

$\Delta f = 232$ Mc

$B = 550$ Gauss

HORN POSITION (cm)

Fig. VII-15. Scattered microwave power vs horn position.

Further evidence for the existence of standing plasma waves along the axis of the tube was obtained by plotting the signal picked up by the strip-line antenna as a function of frequency. The strip line was connected by a coaxial cable to the input of a radio receiver and the output of the receiver is plotted against receiver frequency in Fig. VII-16. These data were taken at a lower magnetic field and a lower beam voltage than the scattering data presented in Fig. VII-15, so comparison with the scattering data is not possible. Figure VII-16 represents the spectrum of oscillations in the plasma and indicates that approximately 25 standing-wave modes are present in the plasma at one time. If we identify each peak with a mode number, n, where n is the number of half-wavelengths between the electron gun and the collector, then we can plot the frequency of the peaks against mode number, as shown in Fig. VII-17. It can be seen that the peaks are all equally spaced in frequency. This is to be expected for
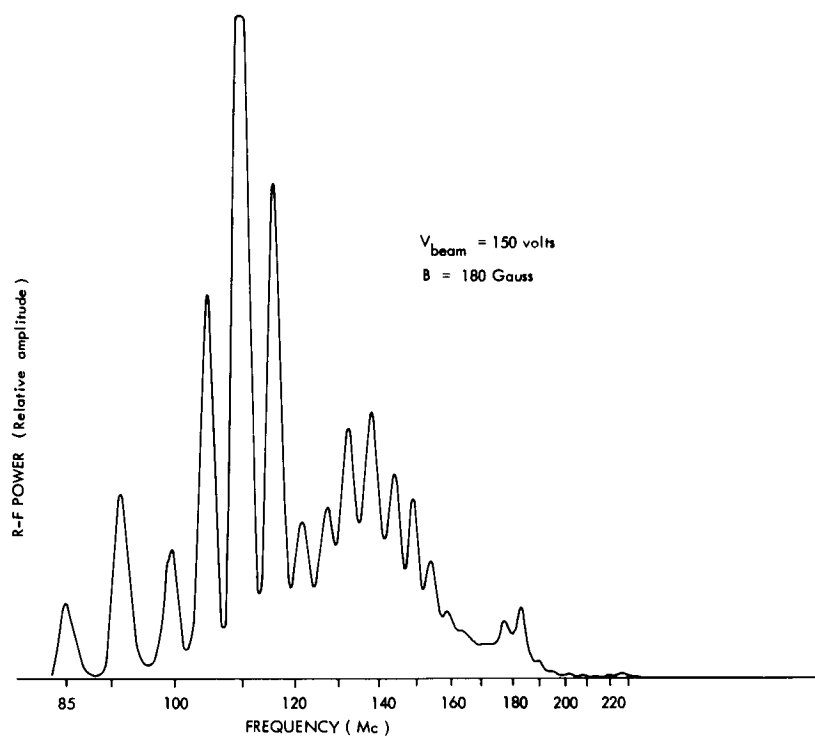
Fig. VII-16. RF power picked up by the strip line from the plasma vs frequency.
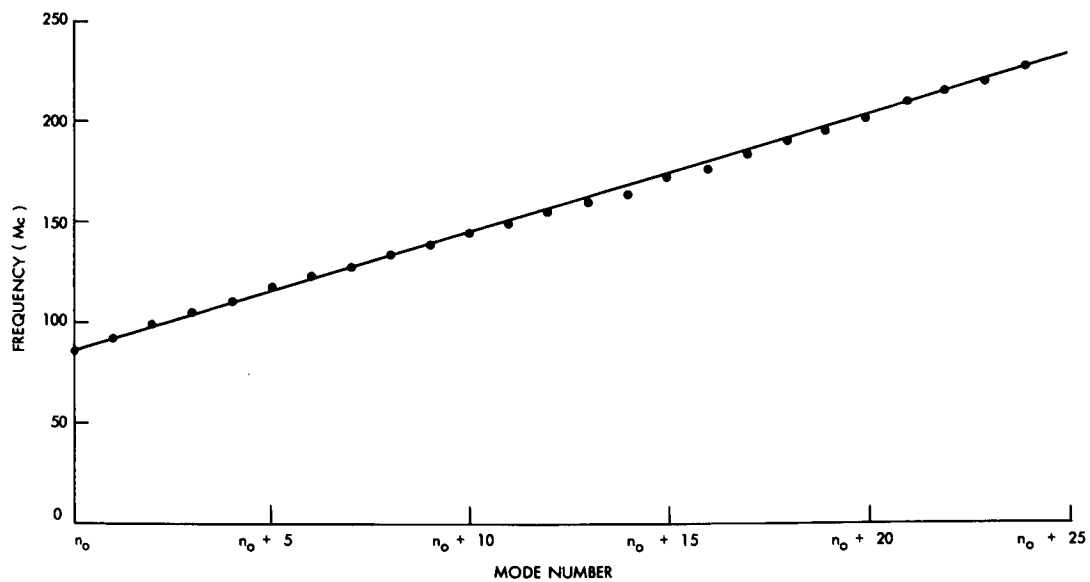
Fig. VII-17. Standing-wave frequency vs mode number.

standing waves that all have the same phase velocity, v_p, and wavelengths given by $D = n\lambda_n/2$, where D is the distance between the electron gun and the collector. Then $f_n = v_p/\lambda_n = nv_p/2D$. From Fig. VII-17 we note that the frequency separation f_n/n is 5.92 Mc. D is 42 cm. We can then calculate the phase velocity, $v_p = \lambda_n f_n = 4.96 \times 10^8$ cm/sec, which is approximately 30 per cent less than the total beam velocity, $v_{beam} = 7.3 \times 10^8$ cm/sec. This is not an unreasonable value for v , the axial beam velocity, since part of the beam velocity is transverse.

A direct measurement of the axial wavelength of one of the standing wave modes shown in Fig. VII-16 was made by moving the strip line along the axis of the tube and
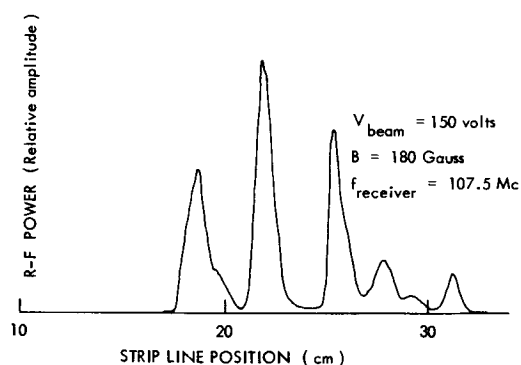
Fig. VII-18. RF power picked up by the strip line from the plasma vs strip-line position.

plotting the amplitude of the signal picked up from the plasma against axial position of the strip line. This is shown in Fig. VII-18. The periodic maxima and minima indicating antinodes and nodes of the standing wave are similar to those shown by the scattering in Fig. VII-15. In Fig. VII-18, however, the frequency of the wave is much lower than that in Fig. VII-15, so the wavelength is longer, and the nodes are more widely spaced. From Fig. VII-18 we get the axial wavelength, $\lambda = 6.30$ cm, and calculate the phase velocity, $v_p = \lambda f = 6.8 \times 10^8$ cm/sec. This is only 7 per cent less than the total beam velocity, which is 7.3×10^8 cm/sec.

It seems quite clear from the data presented here that the microwave scattering is from standing plasma waves excited by the electron beam, since axial movement of both the microwave horns and the strip line showed a standing-wave structure, and the experimentally measured phase velocity was equal to the total beam velocity within 7 per cent. The spectrum of oscillations picked up by the strip line shows that as many as 25 standing-wave modes can be present in the plasma at one time. The assumption that for the nth mode, there are n half-wavelengths between the electron gun and the collector may be in error, since this assumption, together with the experimentally determined frequency difference between modes, implies a phase velocity 30 per cent less than the total beam velocity, which is in disagreement with the phase velocity measured

directly by axial movement of the horns and strip line.

<div align="right">R. L. Kronquist</div>

References

1. R. L. Kronquist, "Microwave Scattering from an Electron-Beam Produced Plasma," Quarterly Progress Report No. 82, Research Laboratory of Electronics, M.I.T., July 15, 1966, pp. 109-114.

VIII. GASEOUS ELECTRONICS[*]

Academic and Research Staff

Prof. S. C. Brown   Prof. J. C. Ingraham   E. M. Mattison
Prof. W. P. Allis   J. J. McCarthy   W. J. Mulligan

Graduate Students

A. J. Cohen   J. E. McClintock   T. T. Wilheit, Jr.
G. A. Garosi   L. D. Pleasance   B. L. Wright

A. COMPARISON OF MEASURED TIME-DEPENDENT ELECTRON VELOCITY DISTRIBUTIONS WITH A THEORETICAL MODEL

In Quarterly Progress Report No. 80 (page 99) a method was discussed for extracting information on the velocity distribution of electrons in a plasma from measurements of the spectrum of emitted microwave radiation. Briefly, in the neighborhood of ω_B, the electron cyclotron frequency of an applied magnetic field, the departure of the effective radiation temperature, $Tr(\omega)$, from a constant value reflects the departure of the electron velocity distribution, $f(v)$, from a Maxwellian. By inversion of an integral relation, $f(v)$ can in principle be determined from $Tr(\omega)$, providing the velocity-dependent electron-atom collision frequency $\nu(v)$, is known. In practice, a shortened expansion of $f(v)$ in Hermite polynomials is adjusted for best fit of the data. The results previously presented illustrated the time dependence of the distribution function early in the afterglow of a weakly ionized argon discharge. At gas pressure ≈ 0.7 Torr and electron density $\approx 10^{10}$ cm^{-3}, it was noted that relaxation to a Maxwellian distribution on a time scale of a few microseconds was accompanied by significant loss of electron energy. Crude analysis seemed to indicate that electron-atom elastic collisions alone could not account for this relatively rapid cooling. It is the purpose of this report to present the results of a more comprehensive attack on that problem.

A theoretical treatment of the behavior of a weakly ionized plasma may be achieved by examination of the Boltzmann equation. We present it here in a form given by Allis.[1]

$$\frac{\partial}{\partial t} f(v, t) = \frac{1}{3} \left(\frac{eE}{m}\right)^2 \frac{1}{v^2} \frac{\partial}{\partial v}\left[\frac{v^2}{\nu}\frac{\partial f}{\partial v}\right]$$

$$+ \frac{2m}{M} \frac{1}{v} \frac{\partial}{\partial(v^2)} [v^3 \nu f]$$

$$+ \sum \left[\nu_x(v_x) \frac{v_x}{v} f(v_x) - \nu_x(v) f(v)\right]$$

$$+ n\left(\frac{e^2}{\epsilon_o m}\right)^2 \ln\left(12\pi n \ell_D^3\right)\left\{f^2\right.$$

$$+ \left[\frac{1}{v^2}\int_0^V V^2 f dV - \frac{1}{3v^4}\int_0^V V^4 f dV + \frac{2}{3v}\int_v^\infty V f dV\right]\frac{\partial f}{\partial v}$$

$$\left.+ \left[\frac{1}{3v^3}\int_0^V V^4 f dV + \frac{1}{3}\int_v^\infty V f dV\right]\frac{\partial^2 f}{\partial v^2}\right\},$$

where e is the charge of the electron, m is its mass, M is the mass of a neutral atom, E is the applied (DC) electric field, ℓ_D is the Debye length, n is the electron density, and $4\pi \int_o^\infty v^2 f dv = 1$. The first term on the right represents the driving force of the applied electric field; the second, elastic recoil of the electron-atom collisions, and the third reflects the transfer of energetic electrons to low velocities through inelastic impacts. Here, $\nu_x(v)$ is the frequency of excitation of an atomic level at an energy eV_x, and v_x is given by $v_x^2 = v^2 + 2eV_x/m$. The last term in brackets describes electron-electron interaction. It is an approximation based on a Fokker-Planck treatment of the fluctuating particle fields. These are the four mechanisms that are thought to be dominant under the conditions of the experiment.

The analysis of the expression above proceeds in two steps. First, the steady-state solution $\left(\frac{\partial f}{\partial t} = 0\right)$ is determined for given values of pressure, density, and electric field ($E \neq 0$). The calculation is performed on a digital computer using a method similar to that employed by Dreicer[2] in his analysis of discharges in hydrogen. Once the steady state has been found, removal of the applied field ($E = 0$) provides the instantaneous time derivative of the distribution function $\left(\frac{\partial f}{\partial t} \neq 0\right)$ at the beginning of the afterglow. The constituent parts of $\partial f/\partial t$ are also examined separately and characteristic rates can be determined for electron-electron, recoil, and inelastic processes.

As an illustration of this method we consider the analysis of the discharge conditions presented in the previous report: argon pressure = 0.723 Torr, electron density = 1.1×10^{10} cm^{-3}. The mean electron energy, U, was found to have a value of 4.56 ev during the pulse and its decay upon removal of the applied field at time t = 0 is illustrated in Fig. VIII-1. The initial rate of energy loss as given by the slope of the straight

ARGON

p = 0.723 torr

n = 1.1 × 10^10 cm^{-3}
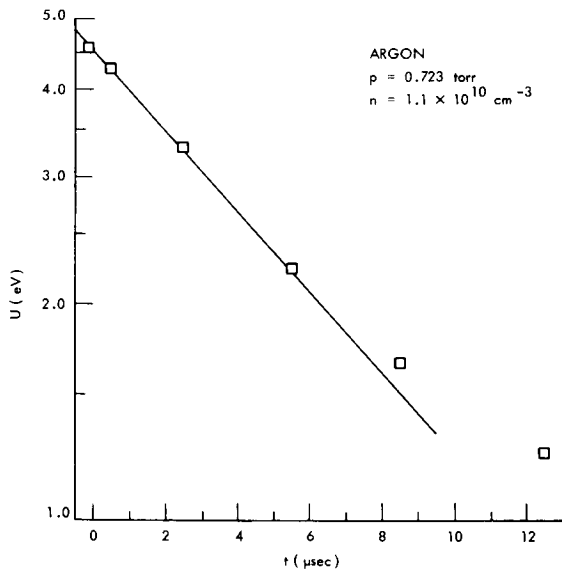
Fig. VIII-1. Afterglow energy decay.

line is 0.59 ev/μsec. The measured argon pressure and electron density were fed directly into the program for computation of the steady-state distribution function. A value of the applied electric field was so chosen that the calculated mean energy would be close to the measured value of 4.56 ev.

A solution obtained with U = 4.54 ev is indicated by the dashed curve in Fig. VIII-2. The solid curve is the velocity distribution inferred from radiation measurements taken during the pulse.

It is also of interest to compare the Tr(ω) peaks generated by these two functions. In Fig. VIII-3, Tr is plotted in terms of a normalized frequency difference

$$\Delta = \left(\frac{m}{2e}\right)^{1/2} \frac{\omega - \omega_B}{100\ p}.$$

The solid and open circles are data points at ω above and below ω_B, respectively. This logarithmic plot tends to emphasize the disparity near $\Delta = 0$, where

EXPERIMENTAL
CALCULATED
MAXWELLIAN

p = 0.72 torr

n = 1.1 × 10^10 cm^{-3}

U = 4.55 eV

$v^2 f(v)$

VELOCITY (square root volts)
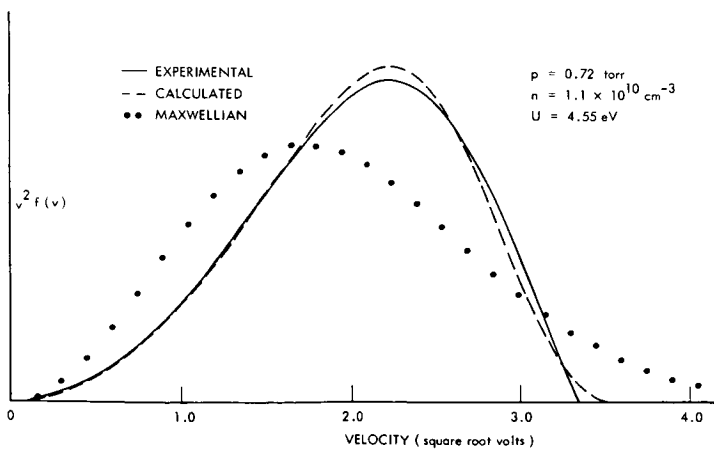
Fig. VIII-2. Comparison of experimentally determined velocity distribution and solution of the Boltzmann equation.
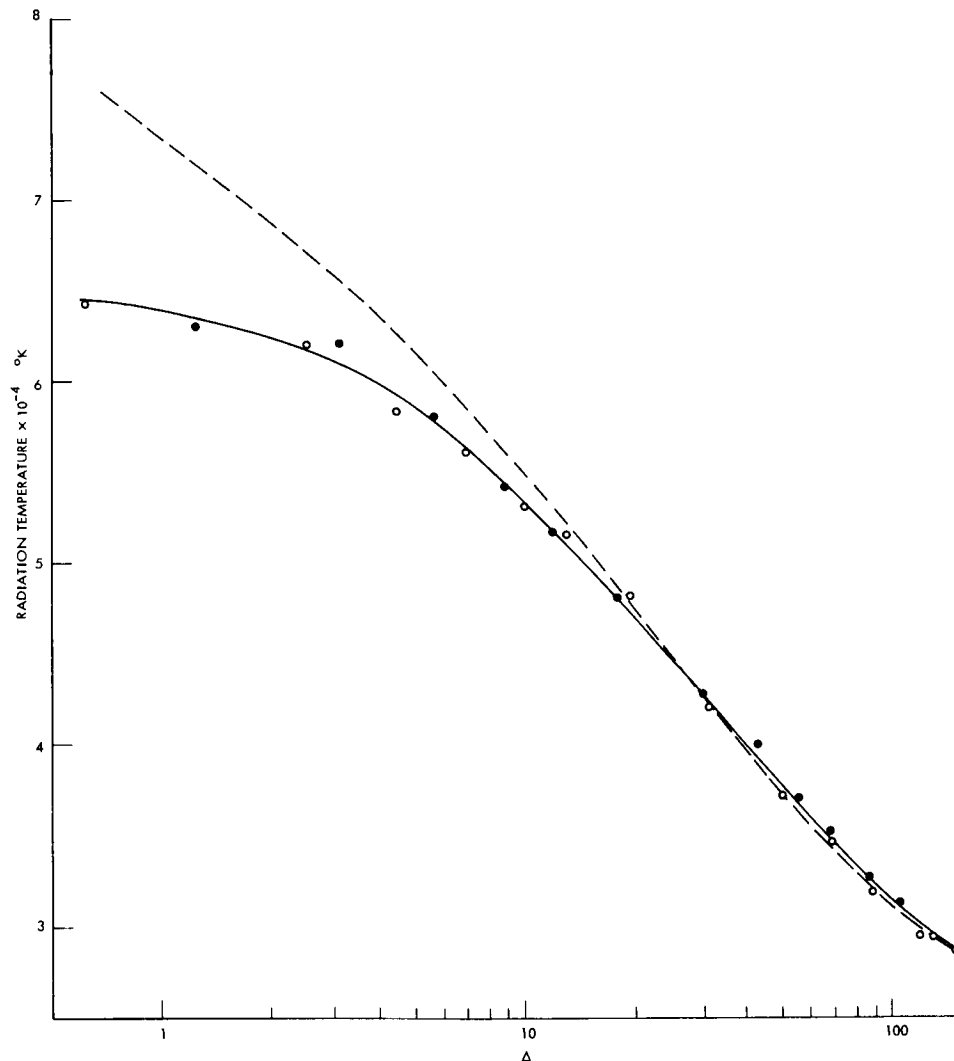
Fig. VIII-3. Comparison of experimental and calculated emission spectra.

$Tr(\omega)$ is a very sensitive function of $f(v)$. Under the conditions of the experiment, Δ is usually certain only within $\pm$ 1.

Figure VIII-4 shows the computed instantaneous time derivative of $v^2 f(v,t)$ at time $t = 0$. The total derivative shown in the top curve offers a rather complex structure, but it is easily understood in terms of the constituent processes presented beneath. The recoil term exhibits a smooth transfer of electrons to lower energies, in contrast with the dumping caused by the first excitation level at 11.5 ev. Integration over velocities reveals that recoil and excitation represent energy loss rates of 0.409 and 0.128 ev/μsec for a combined total of 0.537 ev/μsec. In comparing this figure with the measured value, 0.59 ev/μsec, it should be pointed out that the loss rate is a sensititve function of the
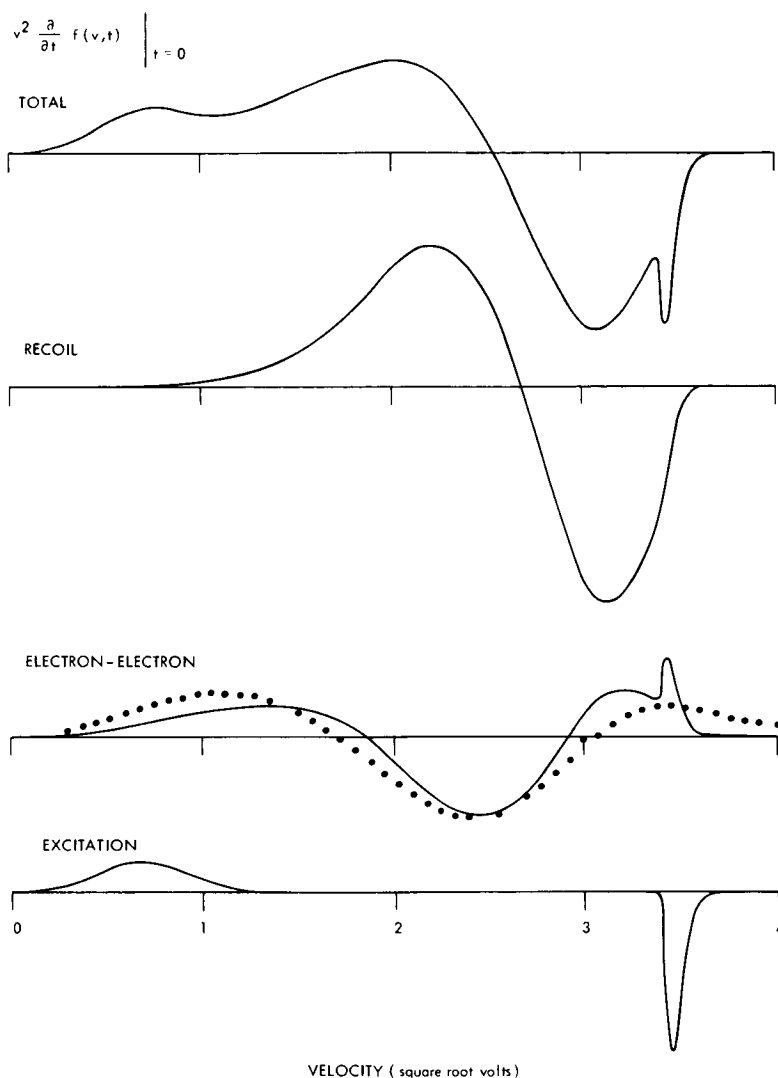
Fig. VIII-4. $v^2 \frac{\partial f}{\partial t}$ at t = 0, showing breakdown into component mechanisms.

mean energy. For instance, a change here in E of 6 per cent would affect U by only 1.75 per cent, but dU/dt would be increased by 10 per cent.

Although the electron-electron interaction term conserves energy, as well as particle number, it clearly participates in the cooling by continuously feeding electrons into the excitation region and thus prevents rapid depopulation. In an attempt to obtain a characteristic electron-electron relaxation time, we have shown for comparison with $v^2 \frac{\partial f}{\partial t}_{ee}$ the dotted curve given by $v^2(f_M - f)/\tau$, where f_M is a Maxwellian distribution with the same mean energy. τ, which represents a time scale for Maxwellization, is chosen for the best fit. In this case $\tau = 4.6$ μsec, which may be compared with

Spitzer's[3] slowing-down time for a test electron of energy U; t_s = 13.8 μsec. A rough estimate of the experimental rate of Maxwellization suggests a figure between 5 and 6μsec.

In conclusion, the various comparisons cited above between observed and expected behavior of the electron velocity distribution seem to substantiate both the experimental method and the theoretical model.

B. L. Wright

References

1. W. P. Allis "Motions of Ions and Electrons," Technical Report 299, Research Laboratory of Electronics, M.I.T., June 13, 1956.

2. H. Dreicer, Phys. Rev. 117, 343 (1960).

3. L. Spitzer, Physics of Fully Ionized Gases, Interscience Publishers, Inc., New York, 1956).

IX. PLASMAS AND CONTROLLED NUCLEAR FUSION

A. Active Plasma Systems*

N67 14630

Academic and Research Staff

Prof. L. D. Smullin	Prof. G. D. Bernard	Prof. M. A. Lieberman
Prof. A. Bers	Prof. W. D. Getty	Prof. J. G. Siambis
	Prof. H. A. Haus	

Graduate Students

R. R. Bartsch	B. R. Kusse	H. M. Schneider
S. R. J. Brueck	J. A. Mangano	M. G. Smith
J. A. Davis	R. R. Parker	R. E. Tremain, Jr.
F. N. Herba	D. M. Perozek	R. N. Wallace

1. ELECTRON BEAM EXCITATION OF ION OSCILLATIONS IN AN ECRD PLASMA

A thesis, entitled "Ion Oscillations Excited by Electron Beam-Plasma Interaction," was submitted by Michael A. Lieberman to the Department of Electrical Engineering, M. I. T. , June 1966, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

A. Bers

2. BEAM-PLASMA DISCHARGE: SYSTEM D

Operation with "Asco" Valve Gas Injection

The plasma density as a function of time has been simultaneously measured by the TM_{010} cavity-mode shift technique and by 8-mm Fabry-Perot interferometry. The density decay (for the minimum gas pulse necessary for a discharge) is shown in Fig. IX-1. The density has been interpreted from the measurements by assuming that the density profile varies as

$$n(r) = n_e J_o(2.405 r/r_{cavity}). \tag{1}$$

The density decay as measured by the Fabry-Perot interferometer in shown in Fig. IX-2 for several different peak pressures (as measured by the swing of the Veeco gauge).

The x-ray spectrum of the discharge above ~130 keV has been measured during various time intervals to determine the shift of the "temperature" of the hot electrons with time. The normalized spectra are shown in Fig. IX-3.

Fig. IX-1.

Density decay for operation with Asco valve.
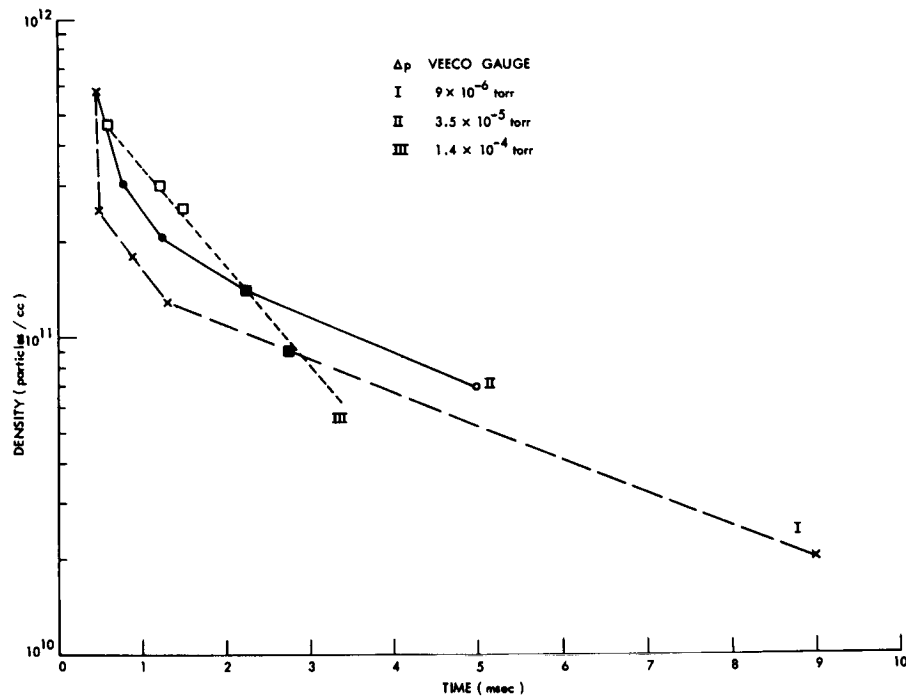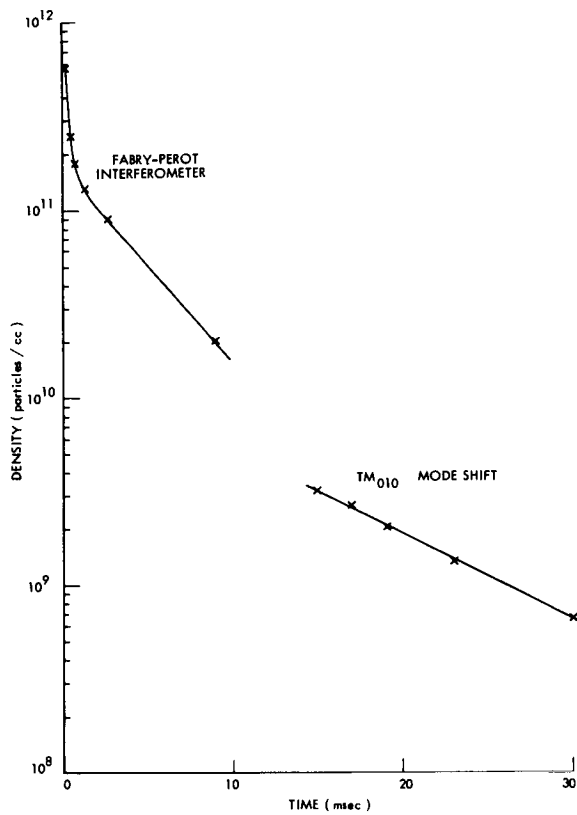
Fig. IX-2. Fabry-Perot interferometer density measurement for operation with Asco valve.
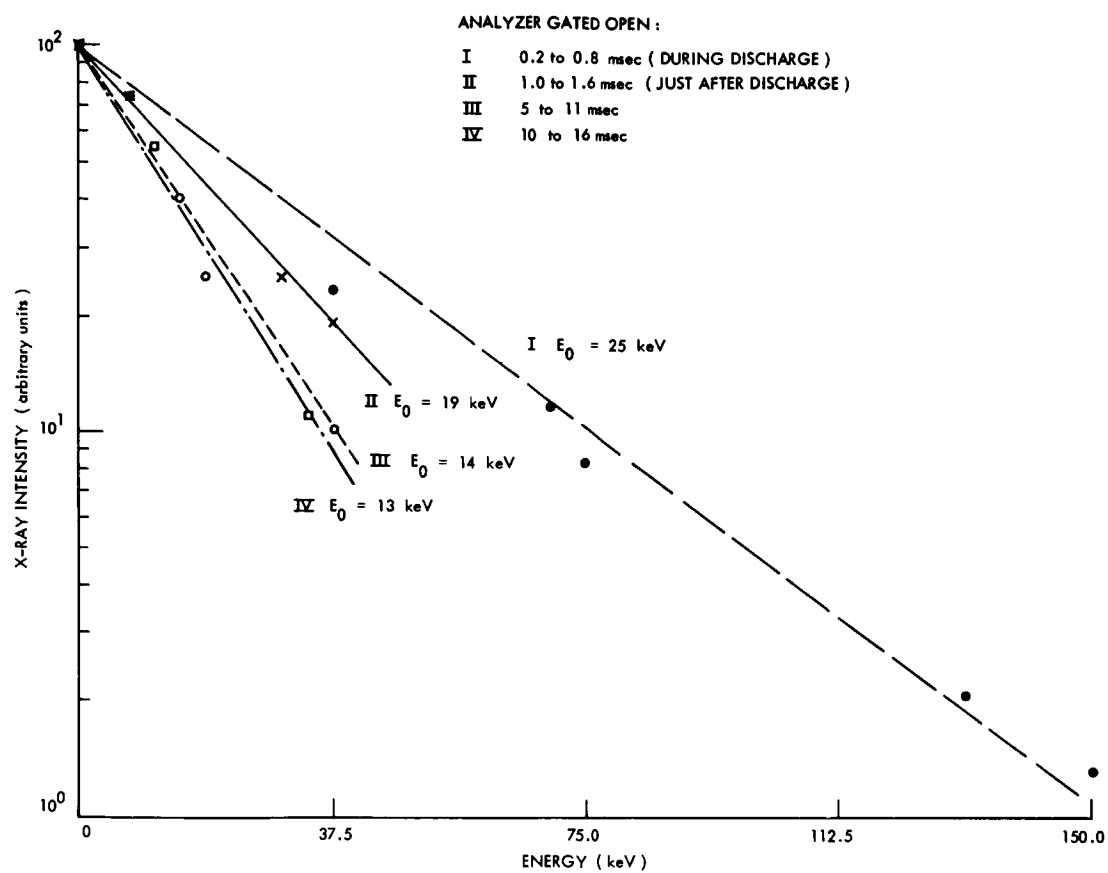
Fig. IX-3. X-ray intensity from discharge.

Operation with Marshall Valve Gas Injection

An electronic drive has been installed on the Marshall valve described previously.[1] The pressure transient as measured by a nude 6AH6 valve is shown in Fig. IX-4. The 6AH6 was located ~2 inches in front of the collector. Similar measurements made for the Asco valve indicate an improvement of ~10 in the rise time of the gas pressure.

Measurements have been made on the discharge obtained with Marshall valve gas injection. The decay following the beam-plasma discharge is observed to be unstable on approximately 50 per cent of the shots.

a. Stable decay: The light output of a stable decay is shown in Fig. IX-5. The peak amplitude of the signal is ~1/10 the peak amplitude of the light signal with Asco valve operation. The uncompensated diamagnetic signal is shown in Fig. IX-6, as an example of the stable decay. The peak amplitude of the diamagnetism is the same ($\sim 10^{14}$ eV/cc) for operation with either valve. The observation that the diamagnetic signal of the stable decay has a hump after the end of the beam pulse was initially interpreted as meaning that the decay time constant of the high-energy electrons was greater than the time constant of flux diffusion. (~20 msec). More complete analysis, by using the method of Parker,[2] gives ~5 msec for the initial decay time constant of the hot-electron density.

The stable density decay has been measured by the techniques used previously on the discharges obtained with the Asco valve. The output of the 4-mm interferometer is shown in Fig. IX-7. Figures IX-8 and IX-9, show the stable density decay for operation with the Marshall valve (interpreted with the plasma radial variation as given by Eq. 1). The light-intensity profile of the discharge was found with the apparatus shown in Fig. IX-10a. The resulting profile corrected for the change of the effective window area with angle is shown in Fig. IX-10b.

b. Unstable decay: The unstable decay is characterized by a sharp drop in the diamagnetic signal. This is accompanied by a sharp drop in the light signal and is illustrated in Fig. IX-11. A burst of x-rays is emitted at the time of the drop of the diamagnetism. This is shown in Fig. IX-12. The spectrum of the emitted radiofrequency has been found to be in the 2-4 kMc range at the time of the instability. The results of suitably filtering the output of a 3/4 inch loop are shown in Fig. IX-13. The X-band output was taken with a waveguide window on the discharge tube. A 2 ft length of S-band waveguide was used as a cutoff filter, and care was always taken to isolate the crystal detector from the system. The observed range of frequencies is in the range of the electron-cyclotron frequency (B_o = 1 to 6 kGauss).

The high-energy x-ray spectrum is shown in Fig. IX-14. The "temperature" is observed to be 51 keV; however, the spectrum analyzer would only detect x-ray energies above ~130 keV. The pulse height analyzer is being repaired and this
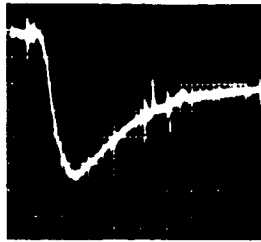
2 msec / cm

Fig. IX-4.   Marshall valve pressure
transient.

0.5 msec / cm

Fig. IX-5.   Marshall valve operation: stable
light decay.
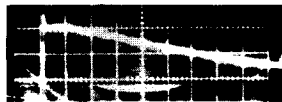
5 msec / cm

Fig. IX-6.   Marshall valve operation:
stable diamagnetic decay.
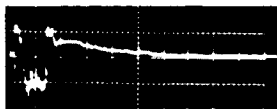
500 μsec / cm

Fig. IX-7.   Marshall valve operation: 4-mm
interferometer density measure-
ment.
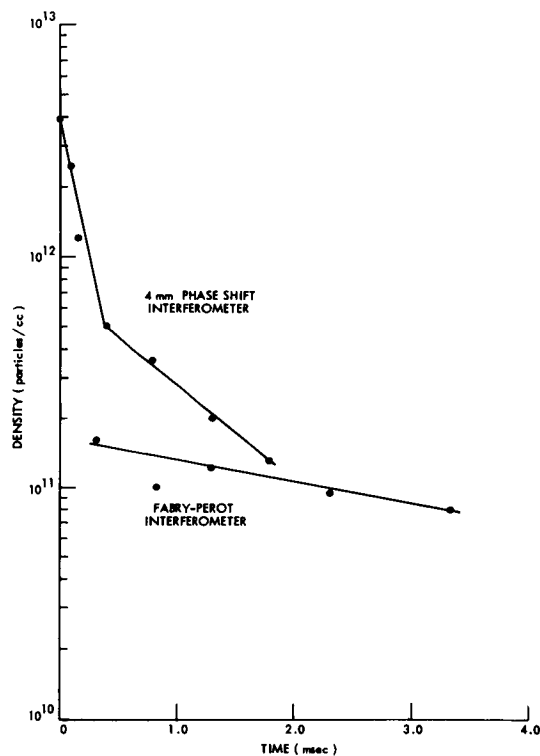
Fig. IX-8. Density decay for Marshall valve operation (high-density range).
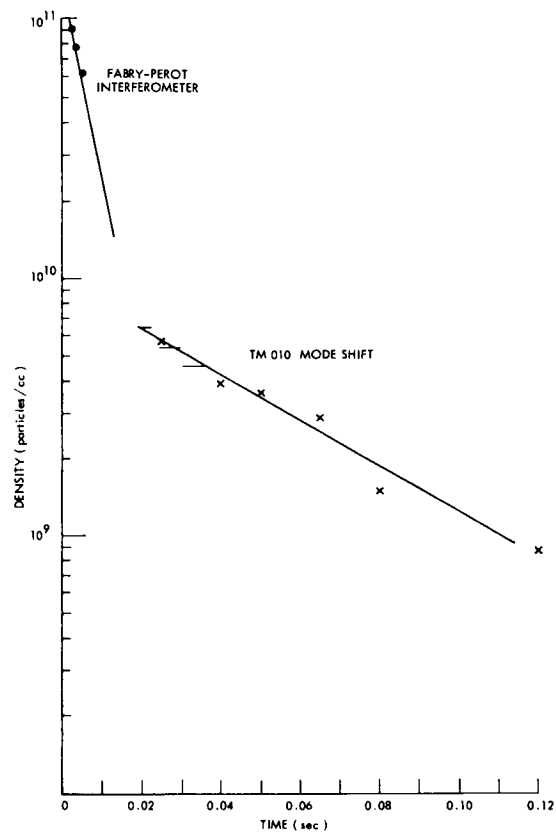
Fig. IX-9. Density decay for Marshall valve operation (low-density range).
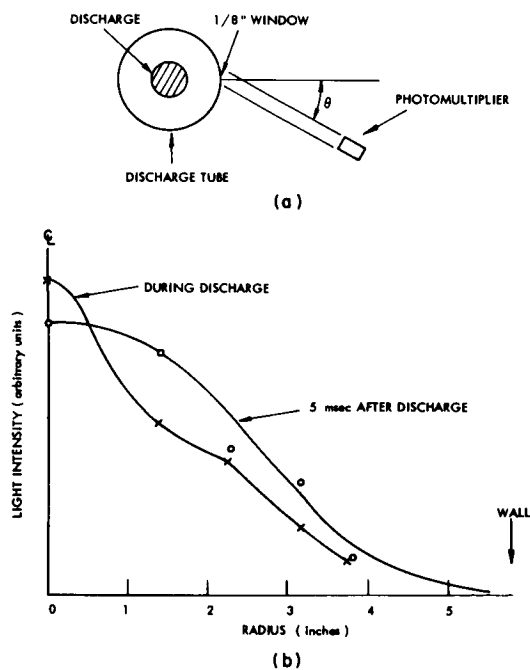
Fig. IX-10. (a) Light-profile apparatus. (b) Light-intensity profile.

LIGHT

DIAMAGNETISM

0.5 msec/cm
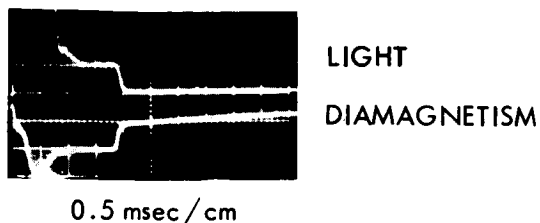
Fig. IX-11. Instability with Marshall valve operation: light and diamagnetism.

X-RAY INTENSITY

DIAMAGNETISM

0.2 msec/cm
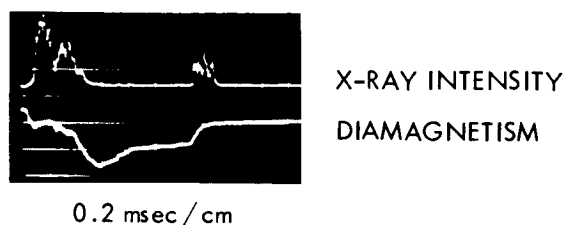
Fig. IX-12. Instability with Marshall valve operation: x-ray and diamagnetism.

400 - 1000 mc

0.5 msec / cm

0 - 1000 mc

0.5 msec/cm

0 - 4000 mc

0.5 msec/cm

X-BAND WAVEGUIDE

1 msec/cm

S-BAND CUTOFF TO 4000 mc

0.5 msec/cm
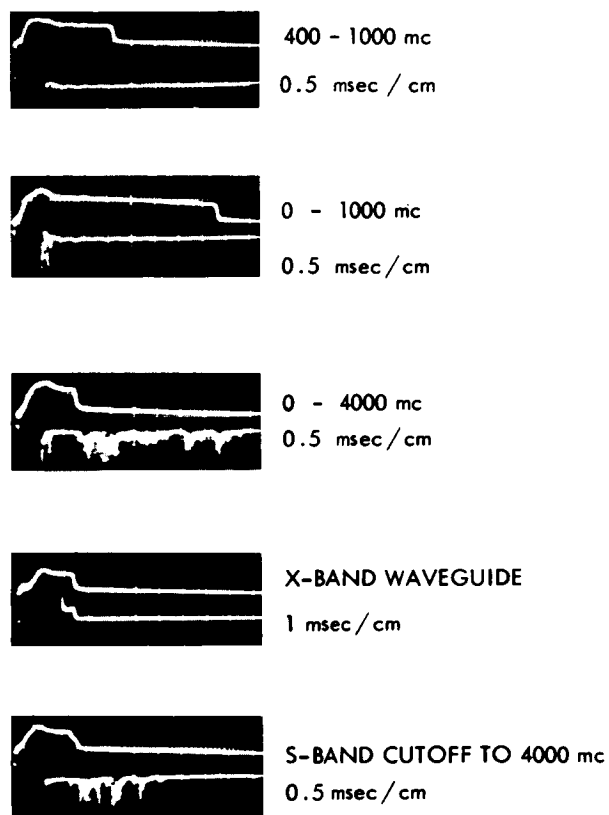
UPPER TRACE : DIAMAGNETISM

LOWER TRACE : RF OUTPUT

Fig. IX-13. Filtered RF output from discharge.

X-RAY INTENSITY
(arbitrary units
logarithmic)

INCREASING ENERGY
37 keV/ division
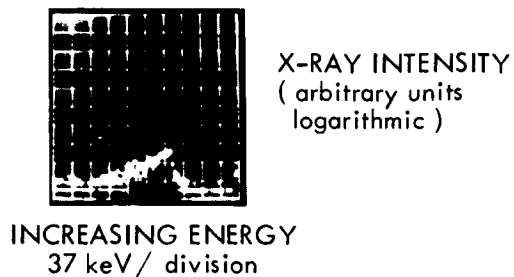
Fig. IX-14. High-energy x-ray spectrum.

measurement will be repeated for a lower energy range (20-60 keV).

Construction of circuits to provide a DC beam of up to 1 kV energy is complete. Experimental investigation of the heating at the ion plasma frequency will be attempted in the afterglow where the plasma consists of hot electrons.[3]

The calibration of the Veeco gauge with a new McLeod gauge is under way at present. Knowledge of the H_2 density will make it possible to compare the decay model described previously with the experimental observations.[4]

The author wishes to acknowledge the use of the facilities of the National Magnet Laboratory for the experiments described here.

R. R. Bartsch

References

1. R. R. Bartsch, Quarterly Progress Report No. 82, Research Laboratory of Electronics, M.I.T., July 15, 1966, pp. 131-137.

2. R. R. Parker, Quarterly Progress Report No. 76, Research Laboratory of Electronics, M.I.T., January 15, 1965, p. 105.

3. M. A. Lieberman, Ph.D. Thesis, Department of Electrical Engineering, M.I.T., June 1966, Chap. IV.

4. R. R. Bartsch, Quarterly Progress Report No. 80, Research Laboratory of Electronics, M.I.T., January 15, 1966, pp. 128-132.

3. HIGH-FREQUENCY ELECTRON-PHONON INTERACTIONS IN A MAGNETIC FIELD

We have previously reported[1] on an analysis of the classical dispersion relation for electrons interacting with acoustic waves in a solid. With parallel applied DC electric and magnetic fields the dispersion relation in the quasi-static approximation ($\bar{q} \parallel \bar{E}$) is

$$K_p + K_e - 1 = 0 \tag{1}$$

$$K_p = \frac{\omega^2 - q^2 s^2}{\omega^2 - q^2 s^2 + Cq^4} \tag{2}$$

$$K_e = 1 + \frac{\omega_p^2}{q^2} \int_{-\infty}^{\infty} dw_\parallel \int_0^\infty 2\pi w_\perp dw_\perp \sum_n \frac{J_n^2(p)\left(\frac{n\omega_c}{w_\perp}\frac{\partial f_{01}}{\partial w_\perp} + q_\parallel \frac{\partial f_{01}}{\partial w_\parallel}\right)}{(\omega - q_\parallel w_\parallel - n\omega_c)} \tag{3}$$

Here, s is the sound velocity, ω_p is the electron plasma frequency, ω_c is the electron-cyclotron frequency, $q_\parallel$ is the wave-number component along B_0, $q_\perp$ is the wave-number component across B_0, $p = q_\perp w_\perp/\omega_c$, C is the coupling constant and electron-lattice collisions and phonon decay have been ignored. For parameters typical of InSb at 77°K and carrier concentrations $\approx 2 \times 10^{14}/cm^3$, Maxwell-Boltzmann statistics are appropriate,

and hence

$$f_{01} = \frac{1}{(2\pi)^{3/2} v_T^3} \exp\left\{-\left[w_\perp^2 + (w_\parallel - v_D)^2\right]\Big/2v_T^2\right\},\tag{4}$$

where v_T is the thermal velocity, and v_D the drift velocity.

Our earlier analysis of the zero magnetic field case was incorrect; the corrected curves of the imaginary part of ω (growth) vs angle of propagation for various electric

$$\omega_s = 5.8 \times 10^{10}/\text{sec}$$
$$v_T = 3 \times 10^7 \text{ cm/sec}$$
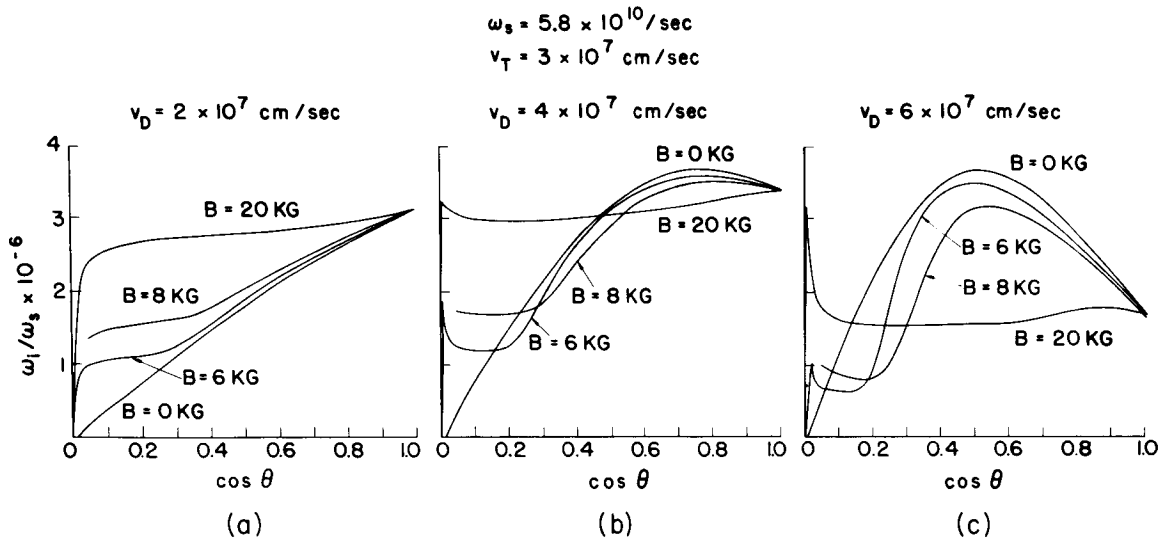
Fig. IX-15. Variation of amplification with direction of propagation. The small cos θ variation has been included only for B = 20 kGauss and B = 6 kGauss.

and magnetic fields appear in Fig. IX-15. As we have previously noted,[1] the growth rate is primarily determined by the imaginary part of the electron dispersion relation.

Zero Magnetic Field

For zero magnetic field the electron dispersion relation reduces to

$$K_e = 1 + \frac{\omega_p^2}{q^2} \int \frac{\overline{q} \cdot (\partial f_{01}/\partial \overline{w})}{(\omega - \overline{q} \cdot \overline{w})} d^3w.$$

The general features of the growth curves in Fig. IX-15 can readily be interpreted on the basis of this equation. The physical mechanism is anti-Landau damping. Figure IX-16 is a plot of the distribution function f_{01} along the direction of propagation. The

anti-Landau damping is expected to be greatest when the electrons near the point of inflection in f_{01} at $w_q = v_D \cos \theta - v_T$ are moving at the wave phase velocity, or

(a)
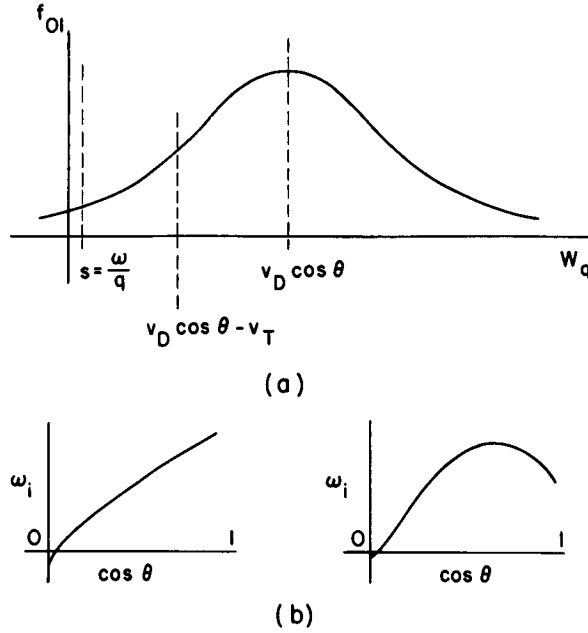
Fig. IX-16. (a) Distribution function f_{01}.

(b) Expected variation in growth rate with respect to angle of propagation for $v_D < v_T$ and $v_D > v_T$.

(b)

$\cos \theta = (s+v_T)/v_D$. It should be noted that for $s + v_T > v_D$ the peak in the growth rate will occur at $\cos \theta = 1$, that is, for propagation along the electric field.

Finite Magnetic Field

When a magnetic field is added the situation becomes considerably more complicated. A new absorption and growth mechanism arises, because of the possibility of cyclotron resonance, and also the anisotropy introduced by the magnetic field invalidates the Landau-damping argument used above for the component of the wave across the magnetic field.

In order to give a physical interpretation of the growth rates, it is convenient to consider the acoustic wave as being drifted rather than the electrons. The dispersion relation, $\omega^2 = \left(q_\perp^2 + q_\parallel^2\right)s^2$, of the acoustic wave is shown in Fig. IX-17. The dotted line on this curve represents the drift velocity. Subject to the constraint of a fixed ω and a fixed $\tan \theta = \dfrac{q_\perp}{q_\parallel}$, it can be shown that the wave becomes negative energy for $\cos \theta > \dfrac{\omega}{qv_D} = \dfrac{s}{v_D}$, and the electrons can be thought of as providing the resistive medium to drive the instability.

Consider, first, the $\dfrac{\partial f_{01}}{\partial w_\parallel}$ term in Eq. 3. The mechanism represented by this term is associated with variations in velocity along the magnetic field and the Landau-damping

argument given above can be applied, with the added feature that now resonances are possible at all cyclotron harmonics. The Bessel functions, whose argument is the ratio
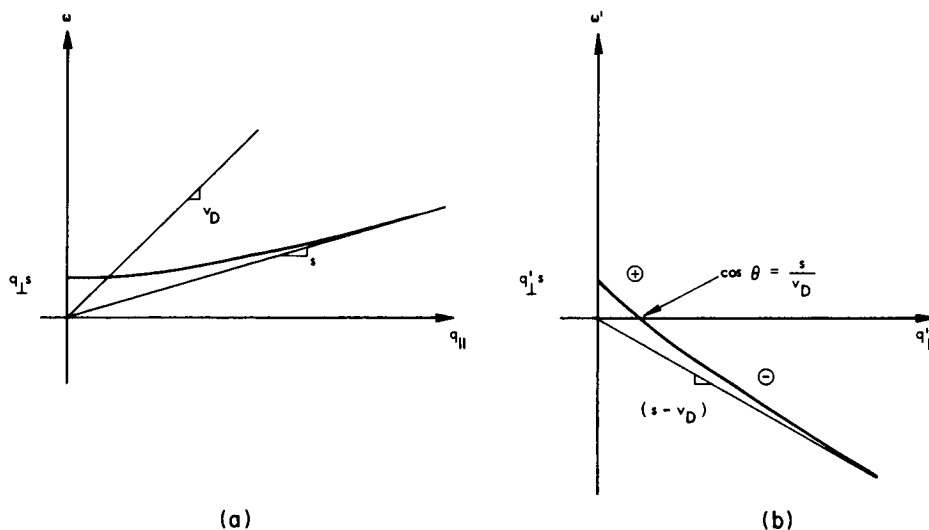
Fig. IX-17. Acoustic-wave dispersion diagram ($q_\perp$ is the parameter).
(a) Laboratory frame. (b) Electron frame with velocity v_D.

of 2π times the Larmor radius of the electrons to the wavelength in the perpendicular direction, give a measure of how the electrons move in achieving this resonance. For example, if an electron has a very large $w_\perp$ so that in the course of one cyclotron period it moves across and back, say, 5 or 6 perpendicular wavelengths, then the net affect would be expected to be small.

There is another mechanism which is given by the $\dfrac{\partial f_{01}}{\partial w_\perp}$ term in Eq. 3. This effect can be explained physically by transforming to a frame moving with the electrons that are in cyclotron resonance with the wave. In this frame, again, with the Bessel function giving some measure of the strength of the interaction, the electrons see a coherent force throughtout their orbit, and hence may either give up energy to the wave or take energy from the wave, their action depending upon their initial phase relationship. As there are more particles with smaller orbits, however, the net effect is to take energy from the wave in this coordinate system.

In order to determine the sign of the interaction in the laboratory frame, it is only necessary to transform the acoustic wave to the moving frame and determine whether or not it is negative energy. For this transformation,

$$\omega^1 = \omega - q_\| w_\| = \omega - q_\| \left(\frac{\omega - n\omega_c}{q_\|} \right)$$

or

$$\omega^1 = n\omega_c,$$

so that for all of the resonances at negative n the wave is negative energy. Since the magnitude of the interaction scales as

$$\frac{f_{01}(w_{||})}{q_{||}} \propto \frac{e^{-(w_{||} - v_D)^2 / 2v_T^2}}{q_{||}}$$

(see Fig. IX-18), the net effect is such as to drive the wave unstable. Again, with
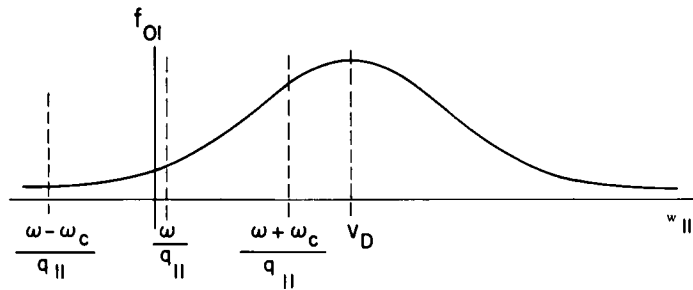
Fig. IX-18. Distribution function along $w_{||}$. For the $n = \pm 1$ terms the wave interacts with the electrons at $w_{||} = \dfrac{\omega \pm \omega_c}{q_{||}}$. The expected variation in growth rate with $v_D < v_T$ and $v_D > v_T$ is then as shown in Fig. IX-15.

reference to Fig. IX-18, it is evident that the greatest gain will be obtained near $\dfrac{\omega + \omega_c}{q_{||}} = v_D$, or $\cos \theta = (\omega + \omega_c)/qv_D$. Thus the effect of the magnetic field is to bring electrons with parallel velocities much greater than the phase velocity of the wave into resonance with the wave.

S. R. J. Brueck, A. Bers

References

1. A. Bers and S. R. J. Brueck, "High-Frequency Electron-Phonon Interactions in a Magnetic Field," Quarterly Progress Report No. 81, Research Laboratory of Electronics, M.I.T., April 15, 1966, pp. 106-111.

IX. PLASMAS AND CONTROLLED NUCLEAR FUSION[*]

B. Applied Plasma Physics Related to Controlled Nuclear Fusion

Academic and Research Staff

Prof. D. J. Rose   Prof. L. M. Lidsky
Prof. T. H. Dupree   Prof. E. P. Gyftopoulos

Graduate Students

R. W. Flynn   C. S. Ribbeck   C. E. Wagner
R. A. Hill   J. C. Woo

1. STUFFED-CUSP PLASMA EXPERIMENT

Construction has been completed on the stuffed-cusp plasma facility. The experimental apparatus has been described previously.[1] The stuffed-cusp magnetic field is generated by a pair of coils operated steady state and a 3-sec pulsed axial current driven by a bank of 12-Volt storage batteries. The required current of 15 KA is obtained by using 11 batteries connected in parallel. Switching is by means of a rack of knife switches driven by a pneumatic cylinder; each battery is switched in parallel to minimize the contact resistance of the circuit breaker. The batteries are then connected in series for recharging.

A pulsed electron beam (1A, 10 kV, 1 msec) is injected through the ring cusp. A beam-plasma discharge is obtained in both cusped and stuffed-cusp geometries in He gas. Diagnostics used thus far have been the monitoring of the RF spectrum with an SP4-A spectrum analyzer, the x-ray output with plastic scintillators, and the light output with a phototube. The results are still preliminary.

Experiments have been proposed to study the effects of the nonzero minimum of the induction on the confinement and stability properties of the cusp.

C. E. Wagner, L. M. Lidsky

References

1. C. E. Wagner and L. M. Lidsky, Stuffed Cusp Plasma Facility, Quarterly Progress Report No. 77, Research Laboratory of Electronics, M.I.T., April 15, 1965.

X. ENERGY CONVERSION RESEARCH*

Academic and Research Staff

Prof. G. A. Brown   Prof. J. L. Kerrebrock   Prof. J. E. McCune
Prof. E. S. Pierson   Prof. M. A. Hoffman
                        Prof. C. C. Oates

Graduate Students

E. K. Levy   R. Decher   W. H. Evers, Jr.
R. J. Thome   G. W. Zeiders, Jr.

A. A LARGE NONEQUILIBRIUM MHD GENERATOR

1. Introduction

It has been clear for some time that nonequilibrium ionization caused by electron heating can be produced in magnetogasdynamic generators, provided the Joule dissipation in the gas can be raised to the level required to elevate the electron temperature.[1,2] Idealized generator calculations predict that it should be quite easy to realize such levels of dissipation. In practice, however, generators operating with the seeded noble gases, which seem optimum for production of electron heating, have, for the most part, failed to yield the expected nonequilibrium ionization. The outstanding exception to this statement is the disc-type Hall generator described by Klepeis and Rosa.[3] The data supporting these conclusions has been summarized by Kerrebrock.[4]

It has been proposed by one of the present authors that the lack of success of the segmented-electrode generators can be attributed to a shorting of the Hall field by a highly conducting layer of gas near the electrode wall.[5] It was proposed that the high conductivity initially results from the concentrated dissipation resulting from the electrode segmentation and is supported by the dissipation caused by the shorted Hall current. The theory appears to explain the behavior of a least two generators, one of which operated at a low Hall parameter, for which the shorting effect was predicted to be small, and the other at a high Hall parameter, for which the effect should be very large.

The success of the disc-type Hall generator is further evidence in support of the theory, since the losses attributable to electrode segmentation are eliminated in this design.

It might then seem that the disc geometry is optimum for nonequilibrium generators. But, since this is a Hall generator, this type suffers from the disadvantage that it _must_ operate at very large Hall parameters in order to yield the levels of efficiency desired for nuclear-MHD power plants. At these large Hall parameters, other losses, such as

those that are due to electrothermal waves[6,7] and steady nonuniformities,[8] may seriously degrade the efficiency.

For these reasons, a systematic attempt has been made to design a segmented-electrode generator which will not exhibit the electrode shorting phenomenon and still be capable of elevation of the electron temperature above the gas stagnation temperature. The design has been based on the results of the theory given elsewhere.[5] It will be described in detail below.

We were fortunate in having available for the study a graphite inert-gas heater capable of a flow of 0.4 kg s^{-1} of helium at 2000°K stagnation temperature and 5 atm stagnation pressure. This makes possible a supersonic flow in which viscous effects are small.

As we shall see, the results of the experiments conducted thus far do not prove the engineering feasibility of segmented electrode nonequilibrium generators. They do serve, however, to indicate some of the critical problems remaining to be solved.

2. Description of the Facility

The hot plasma for use in the experiments is supplied by a large heat-sink facility illustrated in Fig. X-1. This facility is similar in concept to the conventional pebble-bed heaters, but here the core is constructed of 7 separate cylindrical blocks of graphite, 22 inches in diameter, each of which is drilled with approximately 1100 1/4 inch holes. The core is heated electrically by graphite glow bars at an input power of approximately 250 kilowatts. Thermal insulation is provided by graphite felt and graphite fibres. Electrical insulation is provided by boron-nitride spacers.

The test gas is supplied from a manifold of up to 20 gas bottles, and is delivered through two stages of regulators in series, to the inlet at the bottom of the heater. The seed material is injected by a motor-driven stainless-steel syringe into a "bed" of oxygen-free copper bars which have been heated to 1100°K. A subsidiary flow of the flow of the test gas is also passed through the copper bars, and the mixture of vaporized seed material and gas is injected into the plenum chamber at the outlet from the graphite core, where it mixes with the mainstream. The seed is carried from the copper bed to the plenum by a molybdenum tube sheathed with graphite.

No serious problems have been encountered in repeated operation of this facility at 2000°K. This temperature is the approximate limit set by the balance of the available input power against the heat losses.

3. Description of the Channel

a. Over-all Channel Geometry

The MHD channel over-all dimensions were selected to take maximum advantage of the magnet that was used. The magnet has pole face dimensions of 6 × 18 inches and
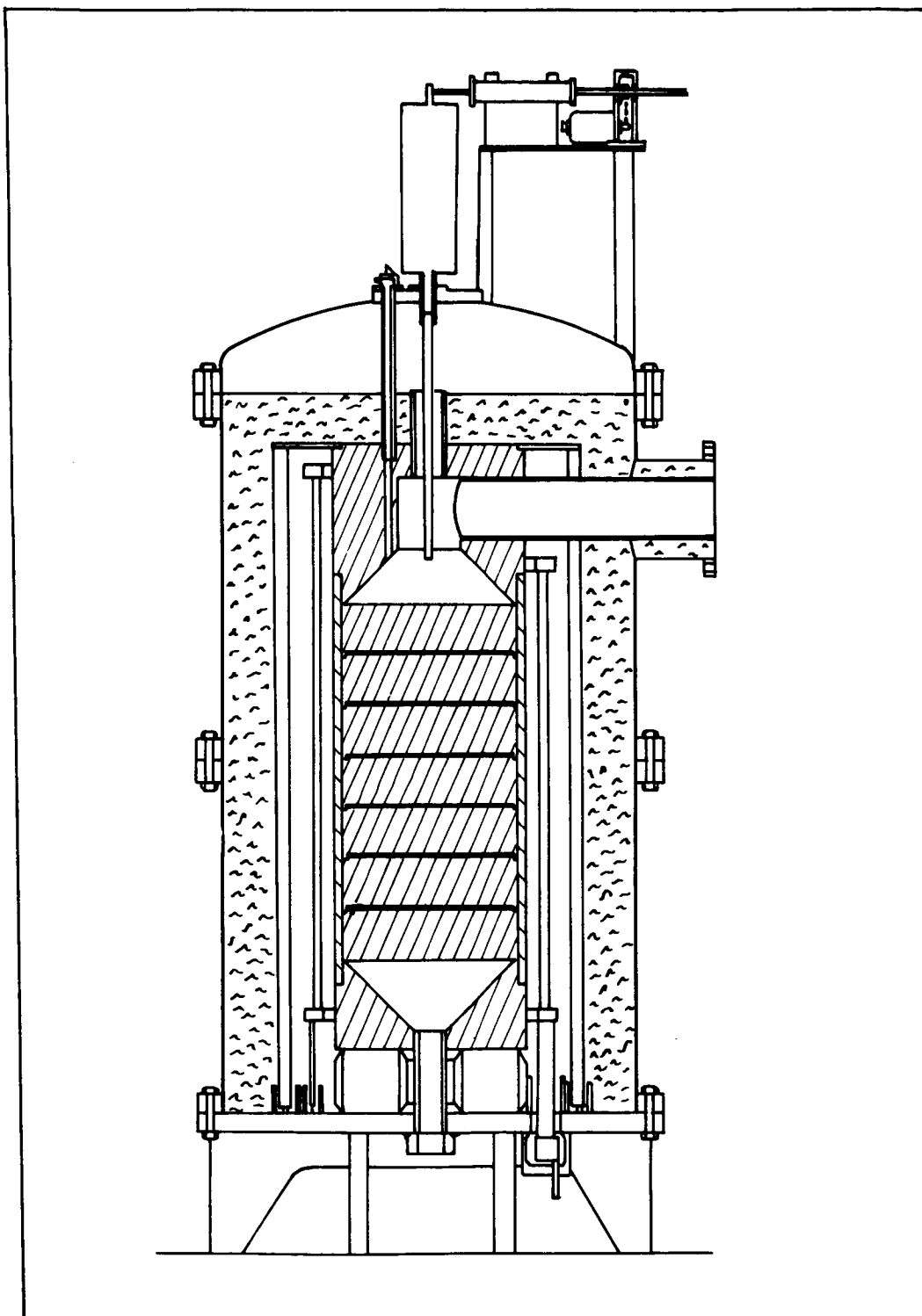
Fig. X-1. Inert-gas heater facility.

produces approximately 1. 3 Tesla at a gap of 2 1/4 inches. As a result, an MHD chan-
nel with a length of 18 inches and an exit height of 6 inches was selected (see Figs. X-2,
X-3, X-4, X-5).

The channel was designed for the following nominal flow conditions: inlet Mach num-
ber of 2. 0 with inlet stagnation temperature of 2000°K, inlet stagnation pressure of 5 atm
and mass flow of helium of 0. 40 kg/sec. This led to the over-all dimensions shown in
Fig. X-2.

A water-cooled copper nozzle of rectangular cross section and constant width was
designed and built to provide the desired inlet Mach number to the channel with a smooth
transition. This nozzle was effectively at ground potential.

A rectangular, constant-area supersonic diffuser was designed to provide good pres-
sure recovery. Great care was exercised in the design to insure electrical isolation of
the MHD channel and plasma from ground. The diffuser walls consist of small copper
bricks or blocks 1 1/2 × 1/2 inches deep mounted on an epoxy plastic support wall. The
blocks have sufficient volume to remain at low temperature throughout the run, since
the run duration is short. Also, each block is electrically and thermally insulated from
adjacent blocks and from the support wall by mica strips, 1/16 inch thick.

A copper subsonic diffuser was attached to the exit of the supersonic diffuser. The
plasma exhausted into a large plenum chamber vented to atmospheric pressure. This
plenum was electrically insulated from ground.

b. Electrode-Wall Design

The electrode geometry and materials were selected on the basis of the theory pre-
viously developed[5] for the behavior of segmented electrodes in a generator where non-
equilibrium ionization of the gas is possible. In order to suppress the electrode-wall
shorting predicted by this theory, it is desirable to use rather coarse segmentation
ratios (i. e. , the ratio of the channel height to electrode pitch) of the order of 5 to 10.
The theory also indicates that it is advantageous to keep the electrode wall cold in order
to suppress the nonequilibrium ionization in the gas layer near that wall.

An electrode pitch of 3/4 inch was selected. This yields a segmentation ratio of 4.7
at the inlet increasing to 8. 0 at the channel exit.

As shown in Fig. X-2, the electrode itself is made of pyrolytic graphite, 5/16 inch
long and 1 1/2 inches wide. The laminae of the pyrolytic graphite lie parallel to the
electrode-wall surface in order to promote heating of the emitting surface of the elec-
trode and reduce heat transfer into the electrode. Each electrode is mounted in a U-
shaped copper yoke to which the external electrical connection is made.

Since the run durations were from 10 to 30 seconds, the heat-sink principle was
employed in designing insulator segments that would remain cold throughout the run.
These consisted of copper blocks of sufficient volume to remain at low temperatures

NOZZLE

MHD CHANNEL

(VIEW WITH TOP INSULATOR WALL REMOVED)

DIFFUSER

SUBSONIC

SUPERSONIC SECTION

18"

15 $\frac{1}{2}$"

5"

$\frac{1}{2}$

5 $\frac{5}{8}$

$\frac{3}{4}$"

ELECTRODE WALL DETAIL

PYROLYTIC GRAPHITE ELECTRODE

COPPER YOKE ELECTRODE HOLDER

EXTERNAL ELECTRODE CONNECTION

COPPER BLOCK

MICA

FIBERGLAS EPOXY PLASTIC SUPPORT WALL

STATION	HEIGHT	WIDTH
A	2 $\frac{3}{8}$"	1 $\frac{1}{2}$"
B	3 $\frac{1}{2}$"	1 $\frac{1}{2}$"
C	6"	1 $\frac{1}{2}$"
D	6"	1 $\frac{1}{2}$"
E	6"	2 $\frac{1}{2}$"
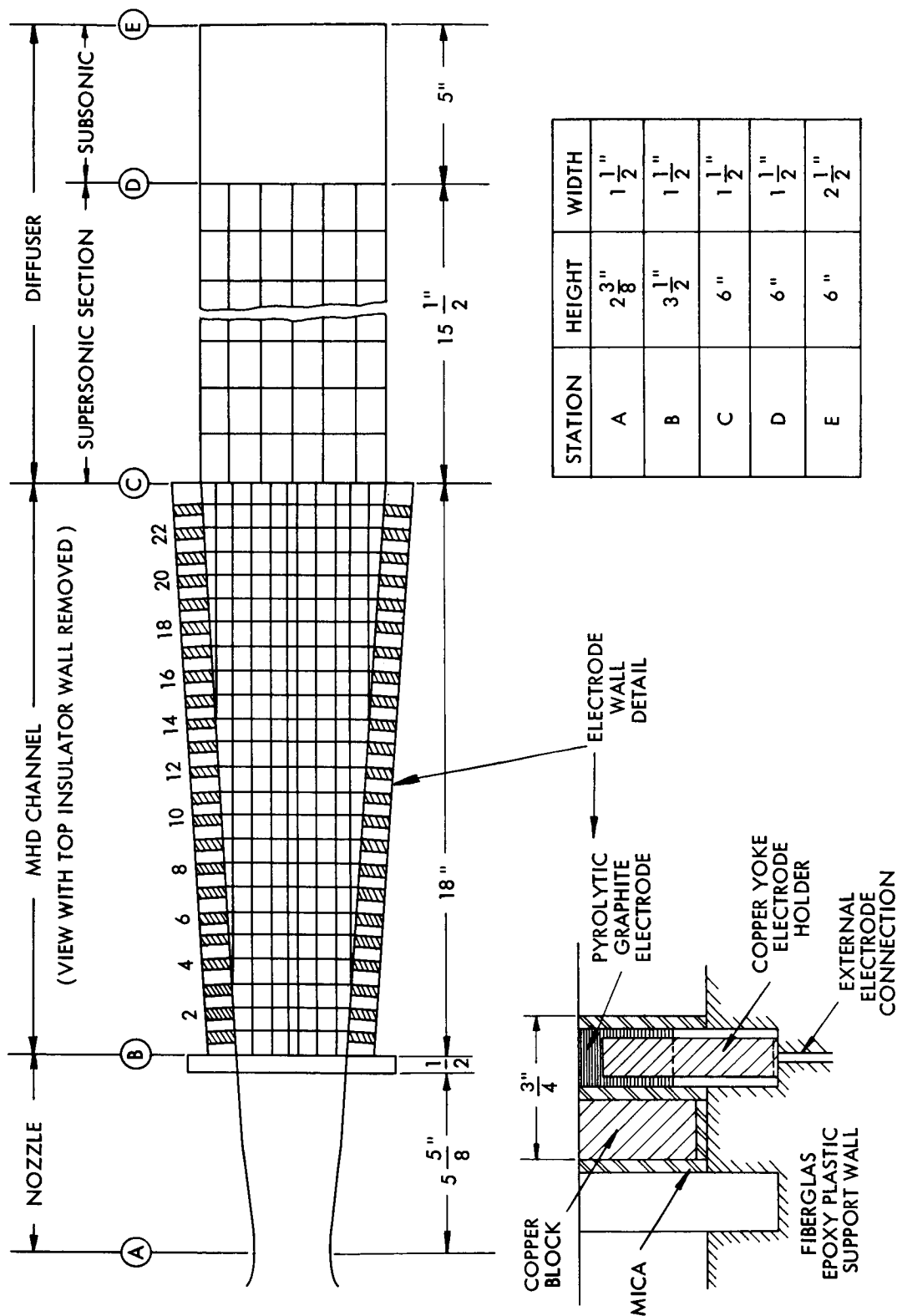
Fig. X-2. MHD channel layout.
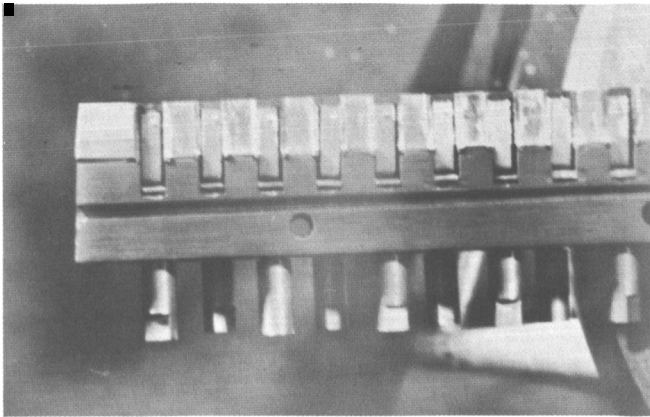
Fig. X-3.

Side view of electrode wall showing first nine segmented electrodes.
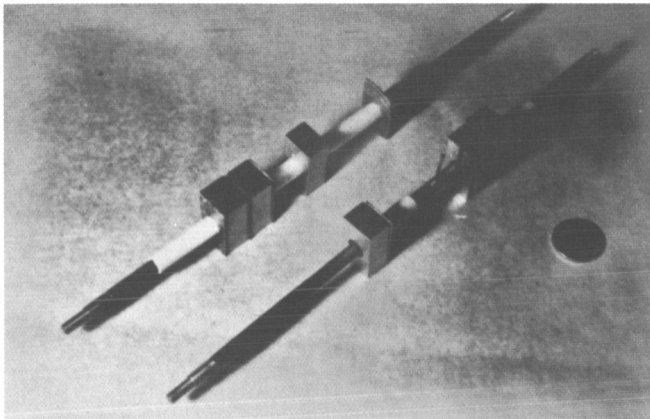
Fig. X-4.

Assembly of Armco iron blocks for insulator wall on steel bolts covered with mica tubes.
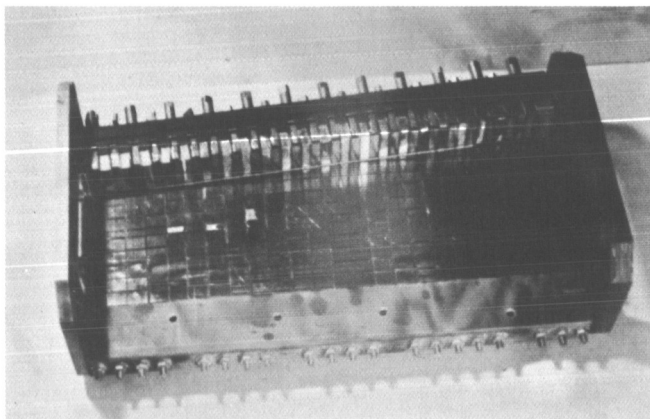
Fig. X-5.

MHD channel partly assembled showing top electrode wall mounted on insulator wall.

(≈200°C) for the expected test durations. These copper blocks were electrically and thermally insulated from the adjacent electrodes and the support wall by mica strips, 1/16 inch thick.

Twenty-three electrodes and insulator segments were assembled on each Fiberglas epoxy plastic-support wall.

c. Insulator-Wall Design

The first insulator wall consisted of sheets of boron nitride, 1/4 inch thick, epoxied to a plastic-support wall. As will be discussed below, insulator wall and/or insulator boundary layer shorting revealed by the results of runs 1 and 2 led to redesign of this wall.

The second insulator wall was assembled from small blocks of Armco magnet iron, each of which was electrically and thermally insulated by thick mica strips, 1/16 inch thick, similarly to the diffuser-wall design. The iron blocks had sufficient depth to remain at relatively low temperature during the run. The use of iron reduced the effective magnet gap to a minimum and increased the magnetic field to 1.4 Tesla.

Each set of blocks lying in the transverse direction was assembled on a long bolt covered with a mica tube for electrical insulation. Holes were drilled along the axis of each bolt and a small pressure tap was inserted into each bolt between two iron blocks. These flush-wall static pressure taps on the insulator wall provided a detailed profile of the static pressure in the flow direction.

The sets of iron blocks were assembled on epoxy plastic-support walls which also provided a pressure-tight seal for the entire insulator wall.

d. Exciter

In order to examine the effect of higher electron density on the starting characteristics of the generator, an external exciter was designed and installed at the entrance of the channel. Several different configurations were used in the various runs. In all cases, a battery bank was used to apply a DC voltage across the channel near the entrance. This battery bank was so connected that it could be inserted or removed when desired during the run at each test condition. In all runs, data were taken at a series of load resistances both with the exciter batteries connected and removed.

In run 2, a 360-volt battery bank was connected across electrode pairs 1, 2, and 3. There was provision for inserting this exciter voltage in series with the load resistor when desired.

In runs 3 to 6, the battery bank was connected only to electrode pair 3. Electrodes 1 and 2 were left open throughout the run. The voltage applied in runs 3, 5, and 6 was 260 volts. In run 4, this voltage was reduced to ~75 volts. In run 6, a 3-ohm resistor was permanently connected across electrode pair 3 throughout the run. As in all other

runs, the exciter batteries could be inserted in series with the load resistor when desired.

The exciter was removed from the electrodes altogether in runs 7 and 8 and connected to the row of insulator wall blocks near electrode pair 2. In run 7, a set of four 12-volt batteries was connected across each of the 7 pairs of insulator wall blocks through a 4-ohm external resistor. These batteries could be applied across the channel so that the current would flow in the direction of the magnetic field. In run 8, a battery bank totaling 260 volts was connected across all 7 pairs of iron blocks with the blocks in parallel.

4. Experimental Results

Eight experimental runs have been conducted thus far, two with the boron-nitride side walls, and six with the iron-peg walls. Of these, the most useful were runs 2, 5, 6, and 7. The principal parameters for these are listed in Table X-1. Helium seeded with cesium was used as working gas in all of the listed runs.

Table X-1. Parameters for experimental runs.

Run No.	Side Wall	P_s (atm)	Cs Mole Fraction	Hall Parameter	T_s (°K)	M	p (atm)	B (Tesla)
2	BN	4.5	.002	3.8	2000	2.1	.50	1.3
5	Iron	4.1	≈0	4.5-6.3	2000	2.1-2.5	.45-.27	1.4
6	Iron	3.1	.003	6.0-8.3	2000	2.1-2.5	.34-.21	1.4
7	Iron	4.5	.002	4.1-5.7	2000	2.1-2.5	.50-.30	1.4

It should be noted at the outset that in none of the runs did the generator produce a significant fraction of the expected power. The data have not been completely analyzed so it will be possible to report only the experimental results, and offer some possible reasons for the poor performance of the generator.

From the pressure distributions (see Fig. X-6), we conclude that the channel operated supersonically, and essentially at the design Mach numbers of 2 at the inlet and 2.5 at the outlet in all runs in which the stagnation pressure was above ~3 atm absolute. As will be noted below, there is a possibility that the magnetogasdynamic effects led to separation on the cathode wall.

The iron-peg-wall channel was operated with a mixture of 70 per cent Argon, 30 per cent Helium, as a Faraday generator in run 3 and as a Hall generator in run 4. In both
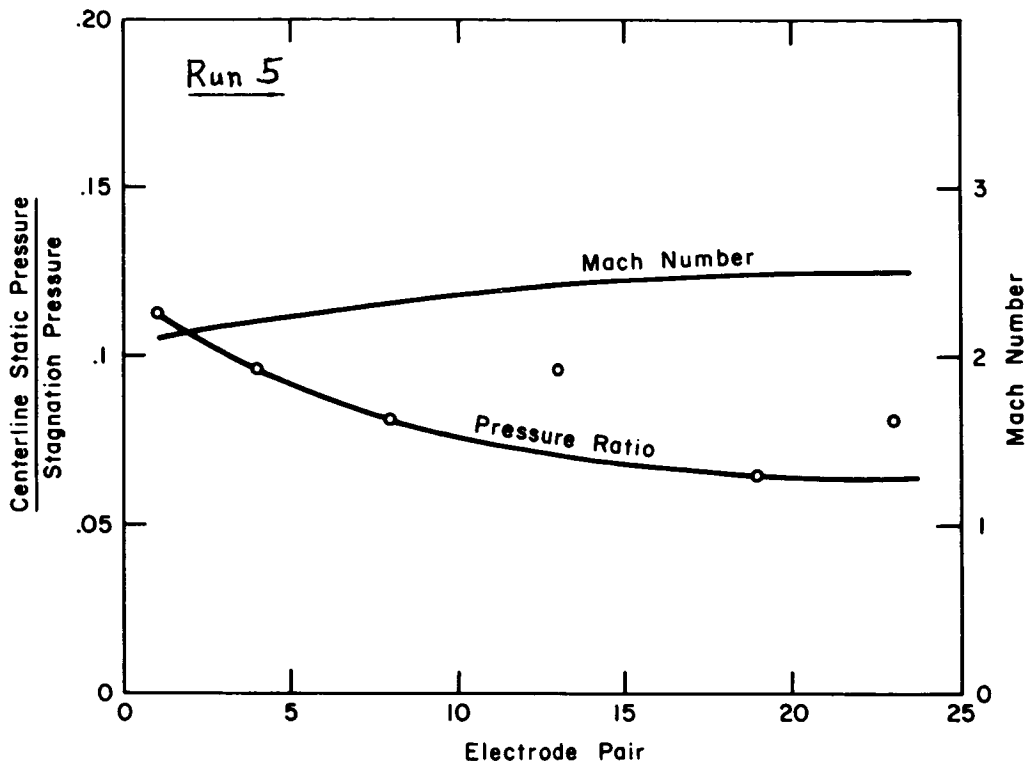
Fig. X-6. Typical static pressure and Mach number variations along the channel.

cases, the open-circuit coltages were very low, and the output negligible. Although the reasons for this are not fully understood, it is probable that the higher Hall parameter and lower electrode temperatures were major contributing factors.

a. Boron-Nitride Side-Wall Runs

The channel with boron-nitride side walls was operated at the test conditions given for run 2 in Table X-1. For these conditions, as for the remaining runs, the conductivity of the plasma should have been approximately 5 mho m^{-1} if the electron mole fraction were frozen at the stagnation value.

An attempt was made to start the generator by applying a potential to electrode pairs 1, 2, and 3 so as to increase the current in these electrode pairs. With the exciter operating, the currents indicated in Fig. X-7 were produced. It will be noted that the current per electrode pair increased down the channel, more rapidly with the 3-ohm load resistance than with the 10-ohm and 25-ohm resistors. From this it may be inferred that the conductivity was increasing down the channel.

The voltage-current characteristics for electrode pairs 5 and 23 are compared (see Fig. X-8) to the characteristic that would be expected for ionization frozen at the stagnation value, and with electrode losses characteristic of the "normal mode."[5] This
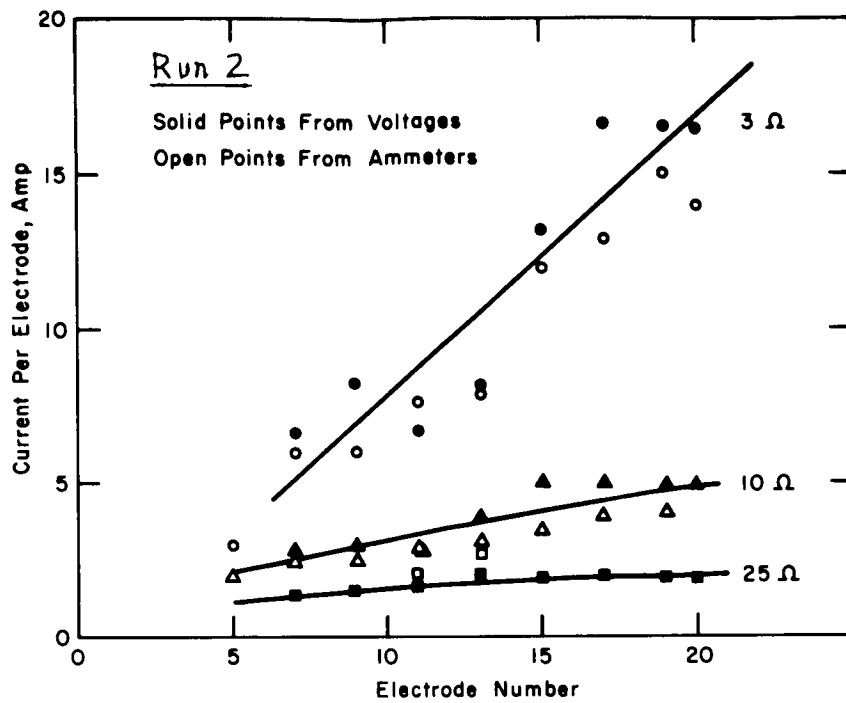
Fig. X-7. Current per electrode pair along the channel for various load resistances.
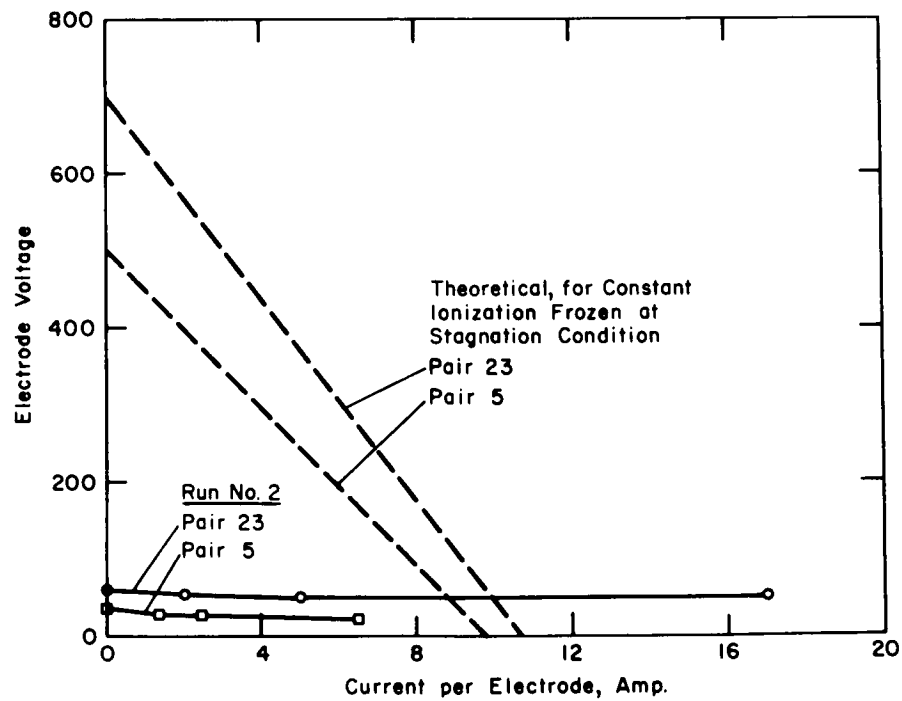
Fig. X-8. Voltage-current characteristics for two represent-ative electrode pairs.

performance is nearly equal to that predicted by the electrode theory of Hurwitz, Kilb, and Sutton,[9] and seems to be a reasonable upper limit to what can be expected if no electron heating occurs.

We see that the open-circuit voltage was very much less than the expected value; however, the current in electrode pair 23 did exceed the frozen ionization value by a substantial factor.

When the exciter was turned off, the currents were negligible.

From the low open-circuit voltage, it was concluded that the generator was probably internally shorted through the boundary layer on the insulating wall. Because of its low thermal conductivity, the boron-nitride side wall ran quite hot, and it was felt that this probably aggravated the shorting.

Accordingly, the iron side walls described above were constructed.

b. Iron Side-Wall Runs

In run 5, the generator was operated with very low seed fraction (because the seeder jammed), both with and without excitation on electrode pair 3. The exciter made very little difference in this run, so the data will be given for points without excitation.

More detailed instrumentation was available for this channel than for the previous one, so that it has been possible to prepare the approximate potential diagrams shown
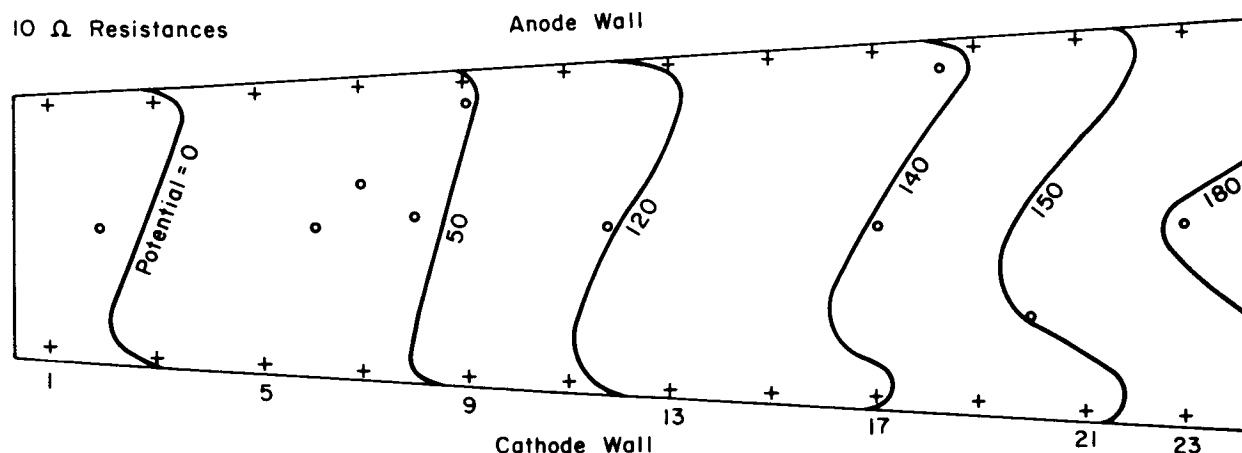
Fig. X-9. Approximate potential diagram for 10-ohm load resistances for run 5.

in Figs. X-8, X-9, and X-10. The crosses indicate points where the voltage was measured at the electrodes, while the circles indicate probes on the insulating wall.

A representative potential diagram from run 5 for the 10-ohm load case is shown in Fig. X-9. The potential pattern is almost symmetric but does show some evidence of a potential "valley" along the cathode wall. This region of low potential tended to
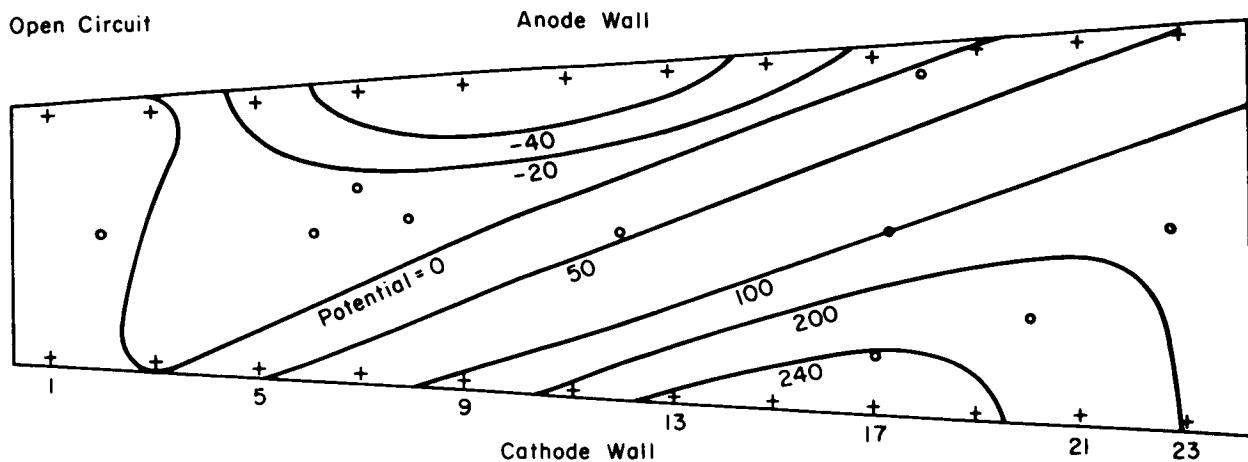
Fig. X-10. Approximate potential diagrams for open-circuit condition for run 6.
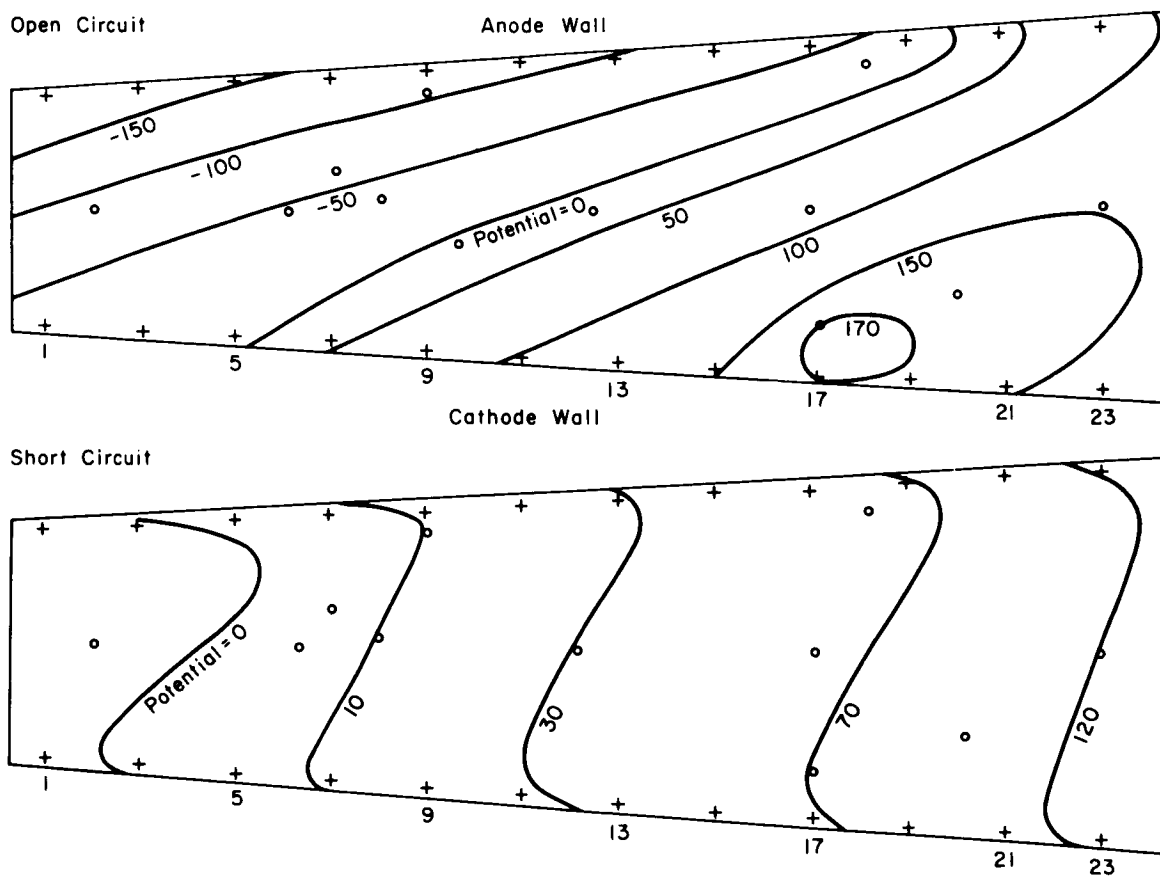
Fig. X-11. Approximate potential diagrams for open-circuit and short-circuit conditions for run 7.

grow at larger load resistances. At lower load resistances it tended to diminish, and essentially disappeared at short circuit.

In run 6, the same channel was operated at a relatively low stagnation pressure, and high seed fraction. The Hall parameter varied between 6.0 at the channel entrance and 8.3 at the exit. A 3-ohm load was maintained on electrode pair 3 for all points. With these conditions, an open-circuit voltage of 365 volts was produced at electrode pair 15. The approximate potential diagram for the open-circuit is shown in condition Fig. X-10. When a load was applied, the channel behaved much as in run 5.

Run 7 was made with almost the same gas conditions as in run 2; data for the cases without excitation are presented. The potential plots for open circuit and short circuit are given in Fig. X-11. Two open-circuit points were taken, one at the beginning of the run and one at the end. At the beginning of the run, the open-circuit voltage reached a maximum of only 75 volts, compared with the 250 volts of Fig. X-11.

Again, the trends are as in runs 5 and 6, with a potential minimum along the cathode wall at open circuit, and large potential drops along both electrodes at short circuit.

5. Discussion of Results

From the potential diagrams, one can construct an approximate current pattern in the channel. We note first that Ohm's law can be written in the form

$$\sigma \vec{E}' = \vec{j} + \beta \frac{\vec{j} \times \vec{B}}{B},$$

where σ is the scalar conductivity, β is the Hall parameter, and $\vec{E}' = \vec{E} + \vec{u} \times \vec{B}$ is the electric field in the gas. Thus, $\vec{j}$ flows at the angle $\tan^{-1} \beta$ to $\vec{E}'$.

Consider first the short-circuit condition for run 7. The directions of $\vec{E}$, $\vec{E}'$, and $\vec{j}$ in the free stream and at the electrode wall are as shown schematically in Fig. X-12a. There is a large Hall current flowing downstream in both locations, while the transverse current flows from anode to cathode, as expected. As the load resistances were decreased, the component of current normal to the electrodes increased, as deduced from rotation of the potential lines so as to be more nearly perpendicular to the axis.

At the surface of the insulator wall, the current flows at the angle $\tan^{-1} \beta$ to the direction of $\vec{E}$, since $\vec{u}$ is small. Thus, the Hall current component reverses and flows upstream. This strongly suggests that the poor performance of the generator in the loaded configuration is due to an axial short of the Hall potential through the boundary layer along the insulating wall.

Next, consider the open-circuit condition for run 7. The current diagram for this case is Fig. X-12b. In the free stream, the current flows more nearly along the axis, as would be expected. At the insulator wall, its Hall component is reversed as before.
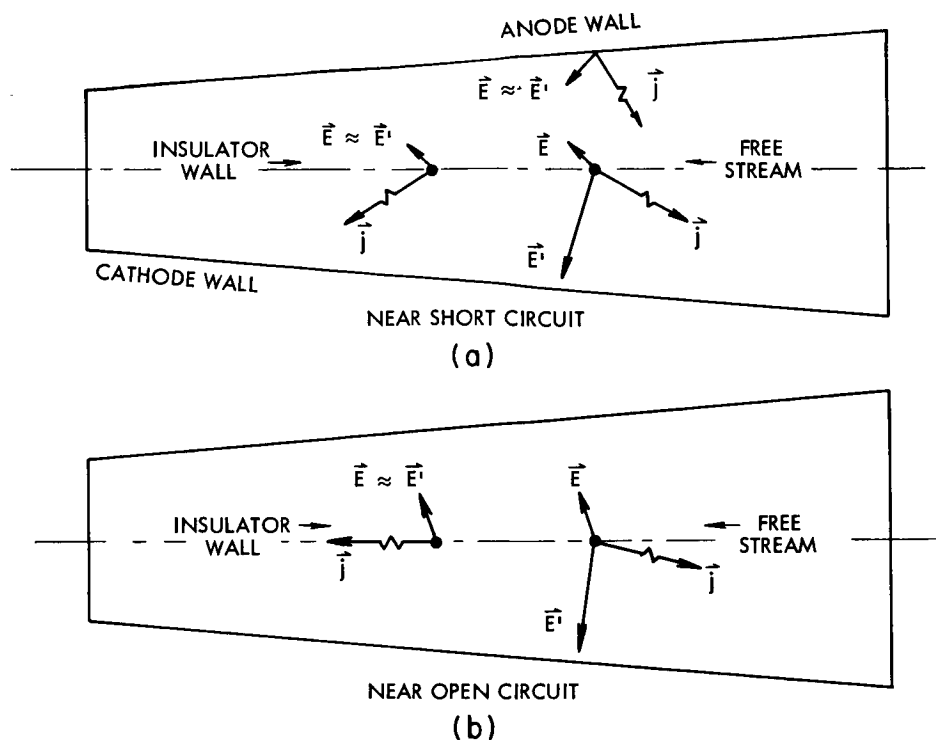
Fig. X-12. Schematic diagram of electric fields and currents for
short-circuit and open-circuit conditions in run 7.

The transverse component may also reverse if the open-circuit voltage is large
enough. To the accuracy that the current diagrams can be drawn, it is diffi-
cult to determine whether the transverse current is positive or negative at this
point.

The low-potential region near the cathode has the appearance of a separated region,
as reported by Louis, Gal, and Blackburn.[10] Such a separation is consistent with a very
large Hall current in the free stream, since such a current would produce a $\vec{j} \times \vec{B}$ force
toward the cathode. Such a large Hall current, however, would have to be supported by
a large transverse current. The return path for this transverse current is not clear, at
present. It may be through the insulator boundary layer.

6. Conclusions

On the basis of these preliminary results, we conclude that the behavior of the insu-
lator boundary layer is a crucial factor in determining the performance of supersonic
inert-gas generators. There seems to be no evidence in the present results that the
electrode shorting phenomenon previously described[5] limited the performance of these
generators.

J. L. Kerrebrock, M. A. Hoffman, G. C. Oates, R. Decher

References

1. J. L. Kerrebrock, "Conduction in Gases with Elevated Electron Temperatures," Engineering Aspects of Magnetohydrodynamics (Columbia University Press, New York, 1962), pp. 327-346.

2. J. L. Kerrebrock and M. A. Hoffman, "Nonequilibrium Ionization Due to Electron Heating: II. Experiments," AIAA J. 2, 1080-1087 (1964).

3. J. Klepeis and R. J. Rosa, "Experimental Studies of Strong Hall Effects and V X B Induced Ionization," Fifth Symposium on Engineering Aspects of Magnetohydrodynamics, Massachusetts Institute of Technology, Cambridge, Massachusetts, April 1964.

4. J. L. Kerrebrock, "Magnetohydrodynamic Generators with Nonequilibrium Ionization," AIAA J. 3, 591-601 (1965).

5. J. L. Kerrebrock, "Segmented Electrode Losses in MHD Generators with Nonequilibrium Ionization-II," Avco-Everett Research Report 201, January 1965.

6. J. L. Kerrebrock, "Nonequilibrium Ionization Due to Electron Heating: I. Theory," AIAA J. 2, 1072-1080 (1964).

7. R. Dethlefsen and J. L. Kerrebrock, "Experimental Investigation of Fluctuations in a Nonequilibrium MHD Plasma," Seventh Symposium on Engineering Aspects of Magnetohydrodynamics, Princeton University, Princeton, New Jersey, 1966, pp. 117-125.

8. R. J. Rosa, "Hall and Ion-Slip Effects in a Nonuniform Gas," Phys. Fluids 5, 1081-1090 (1962).

9. H. Hurwitz Jr., R. W. Kilb, and G. W. Sutton, "Influence of Tensor Conductivity on Current Distribution in an MHD Generator," J. Appl. Phys. 32, 205 (1961).

10. J. F. Louis, G. Gal, and P. R. Blackburn, "Detailed Theoretical and Experimental Study on a Large MHD Generator," AIAA J. 3, 1482-1490 (1965).

COMMUNICATION SCIENCES

AND

ENGINEERING

XI. STATISTICAL COMMUNICATION THEORY*

Academic and Research Staff

Prof. Y. W. Lee   Prof. J. D. Bruce   Prof. D. E. Nelsen
Prof. A. G. Bose   Prof. A. V. Oppenheim   Prof. D. L. Snyder
Prof. H. L. Van Trees†   J. J. Wawzonek

Graduate Students

M. E. Austin†   T. Huang   J. E. Schindall
A. B. Baggeroer†   J. F. Kososki   J. C. Stafford
R. F. Bauer   A. Pitegoff   J. L. Walker
L. D. Collins†   L. R. Poulo   C. J. Weinstein
D. A. Feldman   A. E. Rolland   P. D. Wolfe
                       R. W. Schafer

A. WORK COMPLETED

1. CLASS-D AMPLIFICATION OF RADIOFREQUENCY AMPLITUDE-MODULATED SINE WAVES

This study has been completed by D. A. Feldman. It was submitted as a thesis in partial fulfillment of the requirements for the Degree of Master of Science, Department of Electrical Engineering, M. I. T., August, 1966.

A. G. Bose

2. SPECTRUM ANALYSIS OF THE LOGARITHM OF A FUNCTION

This study has been completed by J. F. Kososki. It was submitted as a thesis in partial fulfillment of the requirements for the Degree of Master of Science, Department of Electrical Engineering, M. I. T., August, 1966.

A. V. Oppenheim

3. AN INVESTIGATION OF TIME JITTER IN NEON BULB THRESHOLD-CROSSING DETECTORS

This study has been completed by J. C. Stafford. It was submitted as a thesis in partial fulfillment of the requirements for the Degree of Master of Science, Department of Electrical Engineering, M. I. T., August, 1966.

D. E. Nelsen

B. OPTIMUM QUANTIZATION OF A TWO-LEVEL SIGNAL AFTER CONTAMINATION BY NOISE

One of the important problems of statistical decision theory is the binary detection problem; that is, the problem of deciding after reception which of two possible signals has been transmitted over a noisy channel. In this paper we shall approach this problem from a point of view different from that traditionally taken.

We begin by considering the message signal s(t) which takes on values s_1 and s_2 with probabilities q and (1-q), respectively. For convenience we assume throughout the paper that $s_1 < s_2$. s(t) is contaminated by an additive, independent noise signal n(t). This results in a received signal x(t),

$$x(t) = s(t) + n(t). \tag{1}$$

The amplitude probability density of the noise signal n(t) is $p_n(\rho)$. Our problem is to determine the two-level quantizer which, when operating on x(t), yields a signal y(t),

$$y(t) = Q[x(t)], \tag{2}$$

that minimizes an appropriate mean value of the quantizer error signal, subject to the constraint that the quantizer output signal can assume only the values s_1 and s_2. That is, an appropriate mean value of the quantizer error signal will be minimized subject to the constraint that the quantizer's representation values are constrained to be s_1 and s_2.

Mathematically, this problem can be formulated in the following manner. We desire that the quantized signal Q[x(t)] be an instantaneous replica of the message portion s(t) of the quantizer-input signal. Generally, we demand more than the quantizer can accomplish. There will be an error,

$$e(t) = s(t) - Q[x(t)]. \tag{3}$$

We shall take an appropriate mean value of e(t) as a measure of how well the quantizer performs with respect to the demands. This measure of the error is given by

$$\mathcal{E} = \int_{-\infty}^{\infty} d\xi \int_{-\infty}^{\infty} d\eta \ \{g[\eta - Q(\xi)] \, p_{x,s}(\xi, \eta)\}. \tag{4}$$

$p_{x,s}(\xi, \eta)$ is the joint amplitude probability density of the quantizer-input signal x and the message signal s (which is also the desired output signal). From our statement of the problem $p_{x,s}(\xi, \eta)$ is

$$p_{x,s}(\xi, \eta) = [q \cdot p_n(\xi - s_1)] \, u_0(\eta - s_1)$$

$$+ [(1-q) \cdot p_n(\xi - s_2)] \, u_0(\eta - s_2). \tag{5}$$

$u_o(t)$ is the unit-impulse function. $g[\eta-Q(\xi)]$ is a function of the error that we call the error-weighting function. $Q(\xi)$ is the quantizer input-output characteristic and is defined in the following way:

$$Q(\xi) = \begin{cases} s_1 & \xi \leqslant x_1 \\ s_2 & \xi > x_1. \end{cases} \tag{6}$$

Substituting Eqs. (5) and (6) into (4) we have for the measure of the error

$$\mathscr{E} = \int_{-\infty}^{x_1} d\xi \int_{-\infty}^{\infty} d\eta \left\{ g(\eta-s_1)[q \cdot p_n(\xi-s_1) u_o(\eta-s_1) + (1-q) \cdot p_n(\xi-s_2) u_o(\eta-s_2)] \right\}$$

$$+ \int_{x_1}^{\infty} d\xi \int_{-\infty}^{\infty} d\eta \left\{ g(\eta-s_2)[q \cdot p_n(\xi-s_1) u_o(\eta-s_1) + (1-q) \cdot p_n(\xi-s_2) u_o(\eta-s_2)] \right\} \tag{7}$$

or, upon integration with respect to η,

$$\mathscr{E} = q \cdot g(0) \int_{-\infty}^{x_1} p_n(\xi-s_1) \, d\xi + (1-q) \cdot g(s_2-s_1) \int_{-\infty}^{x_1} p_n(\xi-s_2) \, d\xi$$

$$+ q \cdot g(s_1-s_2) \int_{-\infty}^{x_1} p_n(\xi-s_1) \, d\xi + (1-q) \cdot g(0) \int_{x_1}^{\infty} p_n(\xi-s_2) \, d\xi. \tag{8}$$

Referring to Eq. (8) we observe that the error $\mathscr{E}$ is a function of only one variable, x_1. We find the value of s_1 which minimizes $\mathscr{E}$ by taking the first partial derivative of $\mathscr{E}$ with respect to x_1 and equating it to zero. Doing this we have

$$0 = q \cdot g(0) \, p_n(x_1-s_1) + (1-q) \cdot g(s_2-s_1) \, p_n(x_1-s_2) - q$$

$$\cdot g(s_1-s_2) \, p_n(x_1-s_1) - (1-q) \cdot g(0) \, p_n(x_1-s_2) \tag{9}$$

or, after algebraic manipulation,

$$\frac{q}{1-q} \frac{g(s_1-s_2) - g(0)}{g(s_2-s_1) - g(0)} = \frac{p_n(x_1-s_2)}{p_n(x_1-s_1)}. \tag{10}$$

The value of x_1 which satisfies (10) minimizes the error $\mathscr{E}$. It should be observed that in most cases of interest the error-weighting function will satisfy the condition

$$g(e) = g(-e). \tag{11}$$

Applying this condition to (10) we have

$$\frac{q}{1-q} = \frac{p_n(x_1-s_2)}{p_n(x_1-s_1)} \qquad (12)$$

which is independent of the error-weighting function.

We now want to compare this result to that obtained using one of the traditional detection theory approaches. The approach which we select for comparison is that of the ideal observer. This point of view maximizes the probability of a correct decision. The x_1 which is optimum from the ideal observer point of view satisfies the equation (see, for example, Wainstein and Zubakov[1])

$$\left. \frac{p_x(\xi|s_2)}{p_x(\xi|s_1)} \right|_{\xi=x_1} = \frac{Pr(s_2)}{Pr(s_1)}. \qquad (13)$$

For this particular problem (13) becomes

$$\frac{p_n(x_1-s_2)}{p_n(x_1-s_2)} = \frac{q}{1-q} \qquad (14)$$

which is identical to (12). Therefore, in this case the detection theory point of view yields a result which is identical to the result obtained by applying optimum quantization techniques.

Consider the following example which indicates how (12) might be solved for x_1. We assume that the noise signal is Gaussian with mean $\bar{n}$ and mean-square value ψ_n. Then,

$$p_n(\rho) = \frac{1}{(2\pi\psi_n)^{1/2}} \exp\left[-\frac{(\rho-\bar{n})^2}{2\psi_n}\right] \qquad (15)$$

Upon substitution of (15) in (12) we have

$$\frac{q}{(1-q)} = \frac{\dfrac{1}{(2\pi\psi_n)^{1/2}} \exp\left\{-\dfrac{[(x_1-s_2)-\bar{n}]^2}{2\psi_n}\right\}}{\dfrac{1}{(2\pi\psi_n)^{1/2}} \exp\left\{-\dfrac{[(x_1-s_1)-\bar{n}]^2}{2\psi_n}\right\}} \qquad (16)$$

After algebraic manipulation this becomes

$$x_1 = \frac{\psi_n}{s_2 - s_1} \ln\left[\frac{q}{1-q}\right] + \frac{s_1 + s_2}{2} + \bar{n}.$$ (17)

which is the desired result.

J. D. Bruce

References

1. L. A. Wainstein and V. D. Zubakov, Extraction of Signals From Noise (Prentice-Hall, Inc., New York, 1962), pp. 150-153.

C. NOISE IN MAGNETIC RECORDING SYSTEMS CAUSED BY THE RECORDING TAPE

1. Introduction

This report summarizes the results of a thesis[1] in which we studied the cause of noise in magnetic recording systems that is due to the recording tape.

Previous investigations have indicated that there are two types of noise which are due to the tape: background noise which limits the smallest signal that can be recorded, and noise that amplitude-modulates the recorded signal and thus broadens its spectrum. This modulation noise, as it is called, is also evidenced by an increase in the noise from the recording system when the tape is uniformly magnetized. It has been felt that these two noise terms are both due to randomness in the magnetic properties of the tape, but no generally accepted and verified theory exists for relating the noise to the physical properties of the tape.

2. Statement of the Problem

The purpose of this study was to relate the observed noise to the physical properties of the recording tape in order to determine its cause.

The procedure was to derive a power density spectrum for the noise flux in the reproduce head in terms of a model incorporating all possible random properties of the tape, which would account for the experimentally observed increase in the noise with uniform tape magnetization. This power density spectrum was then checked through further experiments. These experiments verified the model and made it possible to determine the cause of the noise.

3. Power Density Spectrum of the Flux through the Reproduce Head Because of the Magnetization of the Tape

The relation between the flux through the reproduce head and the magnetization of the tape[1,2] is given by

$$\phi = \int_v \overline{H} \cdot \overline{M} \, dv, \tag{1}$$

where $\overline{M}$ is the magnetic moment density of the tape, $\overline{H}$ is the field set up by the reproduce head for a unit current in the head coil, and v is the volume of the tape.

The magnetic coating of the tape is not continuous, however, but is an assembly of small magnetic particles dispersed in a binder that is coated on a plastic backing.[3]

The flux relation can thus be written

$$\phi(x) = W \sum_i \overline{M}_i \cdot \overline{H}(x-x_i, y_i), \tag{2}$$

under the assumptions that the field function can be adequately represented in just two dimensions,[2] and the variation of the field over the volume of a particle is negligible. W is the width of the tape.

The autocorrelation function of the head flux in terms of the random properties of the magnetic particles[1] is

$$R_\phi(\tau) = W^2 \int_0^\infty k^2 |F(k)|^2 \, e^{-2k(d+T/2)} \frac{\sinh^2\left(k\frac{T}{2}\right)}{\left(k\frac{T}{2}\right)^2} \left\{ \overline{N} E[m^2]\left(k\frac{T}{4}\right) \frac{\sinh\left(k\frac{T}{2}\right)}{\cosh\left(k\frac{T}{2}\right) - 1} \right.$$

$$\left. + 2\overline{N} E^2[\cos\theta] \, E^2[|m|] \sum_{n=1}^\infty E[s_i s_{i+n}] \, \text{Re} \, M_{D_{i,i+n}}(k) \right\} \cos(kv\tau), \tag{3}$$

where $F(k)$ is the two-dimensional transform of the magnetic scalar potential of the head, d is the spacing between the head and tape, T is the thickness of the magnetic coating of the tape, $\overline{N}$ is the average number of particles per unit length of the tape, m is the particle magnetic dipole moment, θ is the angle from the x axis of the projection of the particle on the x-y plane, s is the sign of the particle, $M_{D_{i,i+n}}(k)$ is the characteristic function of the random distance between the ith and the $(i+n)^{th}$ particles, and v is the tape velocity.

The power density spectrum of the flux is

$$S_\phi(\omega) = W^2 \left(\frac{\omega}{v}\right)^2 \left|F\left(\frac{\omega}{v}\right)\right|^2 e^{-2\frac{\omega}{v}(d+T/2)} \frac{\sinh^2\left(\frac{\omega}{v}\frac{T}{2}\right)}{\left(\frac{\omega}{v}\frac{T}{2}\right)^2} \left\{ \overline{N} E[m^2]\left(\frac{\omega}{v}\frac{T}{4}\right) \frac{\sinh\left(\frac{\omega}{v}\frac{T}{2}\right)}{\cosh\left(\frac{\omega}{v}\frac{T}{2}\right) - 1} \right.$$

$$\left. + 2\overline{N} E^2[\cos\theta] \, E^2[|m|] \sum_{n=1}^\infty E[s_i s_{i+n}] \, \text{Re} \, M_{D_{i,i+n}}\left(\frac{\omega}{v}\right) \right\}. \tag{4}$$

Correcting this equation for the transfer characteristic of the head gives the power density spectrum of the magnetization:

$$S_M(\omega) = \overline{N} E[m^2] \left\{ \left(\frac{\omega}{v}\frac{T}{4}\right) \frac{\sinh\left(\frac{\omega}{v}\frac{T}{2}\right)}{\cosh\left(\frac{\omega}{v}\frac{T}{2}\right) - 1} + 2E^2[\cos\theta] \frac{E^2[|m|]}{E[m^2]} \right.$$

$$\left. \sum_{n=1}^{\infty} E[s_i s_{i+n}] \operatorname{Re} M_{D_{i,i+n}}\left(\frac{\omega}{v}\right) \right\}. \tag{5}$$

This power density spectrum can be rewritten

$$S_M(\omega) = \overline{N} E[m^2] \left\{ \left(\frac{\omega}{v}\frac{T}{4}\right) \frac{\sinh\left(\frac{\omega}{v}\frac{T}{2}\right)}{\cosh\left(\frac{\omega}{v}\frac{T}{2}\right) - 1} + 2E^2[\cos\theta] \frac{E^2[|m|]}{E[m^2]}(P_+ - P_-)^2 \sum_{n=1}^{\infty} \operatorname{Re} M_{D_{i,i+n}}\left(\frac{\omega}{v}\right) \right.$$

$$\left. + 8E^2[\cos\theta] \frac{E^2[|m|]}{E[m^2]} \sum_{n=1}^{\infty} \left(P_+ P_- - P_{+-}^{(n)}\right) \operatorname{Re} M_{D_{i,i+n}}\left(\frac{\omega}{\theta}\right) \right\}, \tag{6}$$

where $P_+ = p(s_i = +1)$, $P_- = p(s_i = -1)$, and $P_{+-}^{(n)} = p(s_i = +1, s_{i+n} = -1)$. Note also that $P_+ - P_-$ is proportional to the magnetization of the tape.

Examining Eq. 6, we see that the first term is a noise term independent of the magnetization, the second term contains the signal and can also contain a noise term that is dependent on the magnetization, and the third term is also a noise term dependent on magnetization. These two noise terms, which change with magnetization, are caused in two different ways: the first is due to the distribution of the magnetic particles, and the second to interaction between the particles.

4. Experimental Results

We shall now present the experimental results that allow us to determine the cause of the modulation noise. In Fig. XI-1 we show the power density spectra of the flux for various levels of magnetization from the erased state to the saturated state. These curves show the manner in which the power density increases with magnetization.

To determine the cause of the noise that increases with magnetization, we correct these spectra for the transfer function of the head and compare them with the theoretical spectrum from Eq. 6. We note that if the modulation noise is caused by the particle distribution alone, the shape of the spectrum in the erased state is
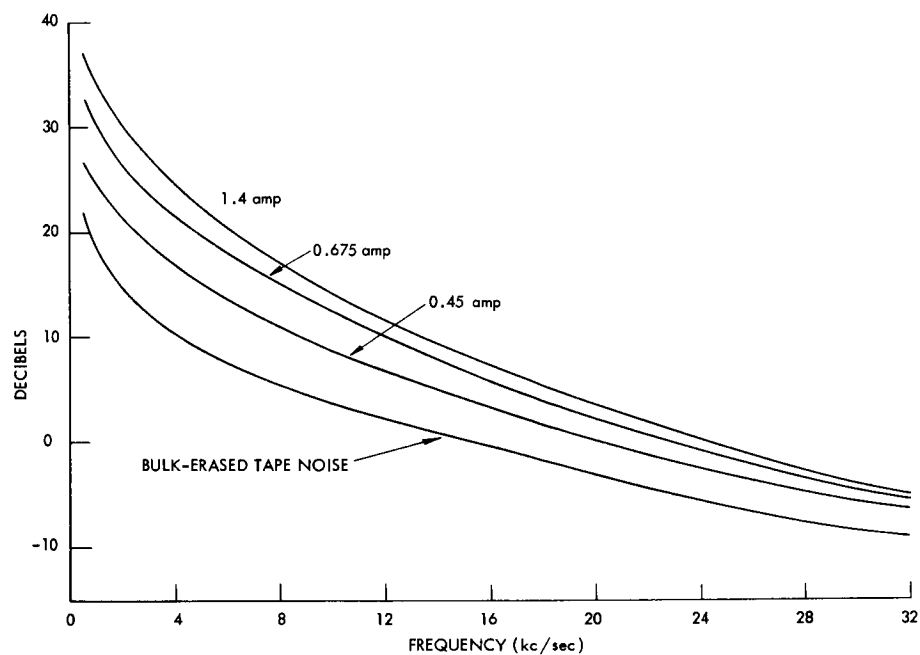
Fig. XI-1. Power density spectrum of noise through the reproduce head as a function of recording current. Tape speed: 7.5 ips.
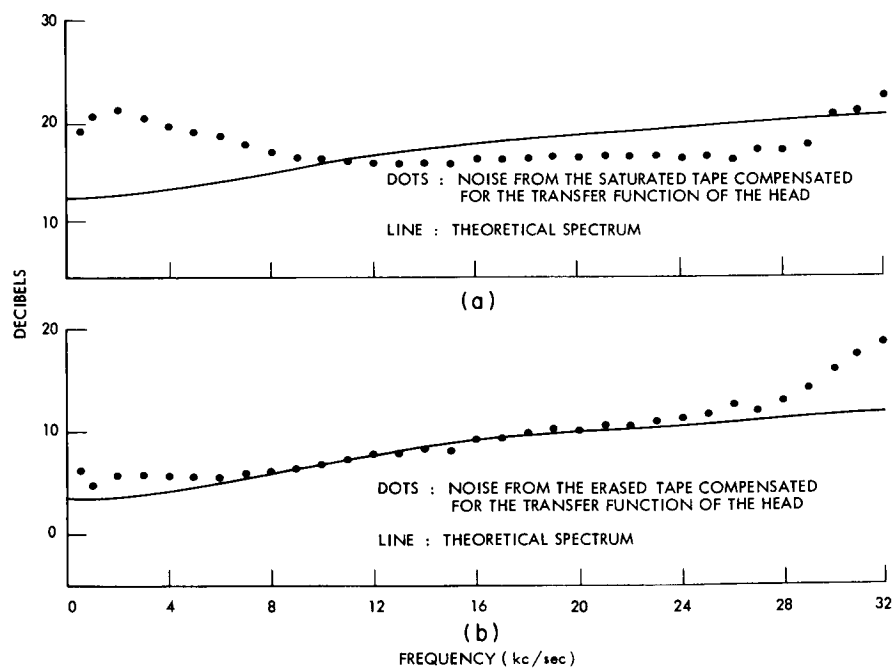
Fig. XI-2. Comparison of measured power density spectra in the (a) saturated and (b) erased states with f(ω).

$$f(\omega) = \left(\frac{\omega}{v}\frac{T}{4}\right) \frac{\sinh\left(\frac{\omega}{v}\frac{T}{2}\right)}{\cosh\left(\frac{\omega}{v}\frac{T}{2}\right) - 1},\tag{7}$$

while if the modulation noise is caused by particle interaction alone the shape of the spectrum in the saturated state will be f(ω).

In Fig. XI-2 we compare the measured power density spectra in the erased and saturated states with f(ω). This shows that the measured noise spectrum from the erased tape most closely matches f(ω). This leads us to the conclusion that the modulation noise is caused by the manner in which the particles are distributed on the tape.

The spectral shape of the noise, which increases with magnetization, is determined by subtracting the noise that is due to the erased tape from the other spectra. The resulting power density spectra for various values of magnetization, corrected for the
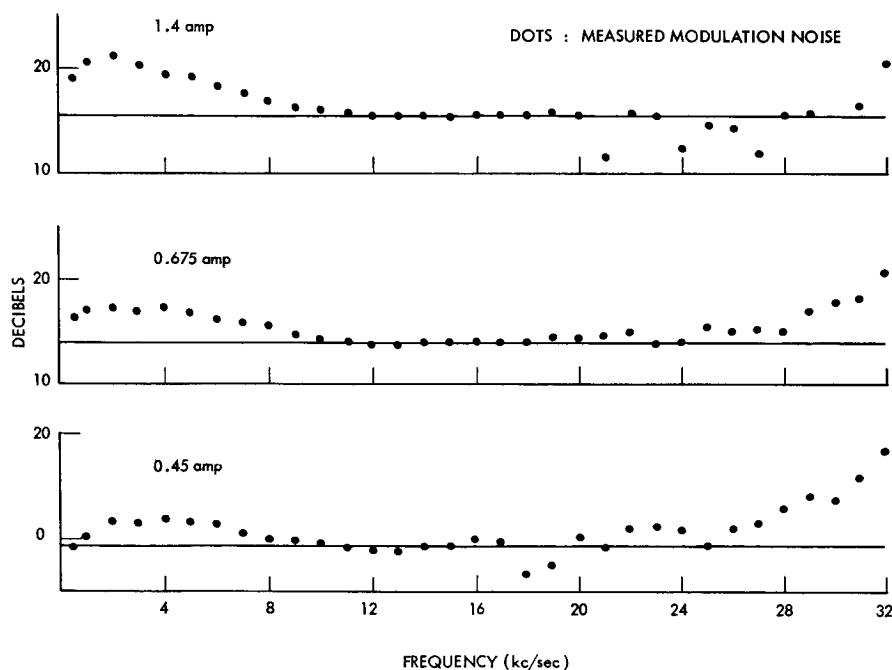
Fig. XI-3. Power density spectra for various values of magnetization corrected for the transfer function of the head.

transfer function of the head, are shown in Fig. XI-3. These plots indicate that the modulation-noise term has a relatively flat spectrum in the frequency range 0-30 kc.

5. Conclusions

We have derived the power density spectrum for the flux through the reproduce head of a magnetic recorder in terms of the random properties of the magnetic particles in

the recording tape. This power density spectrum has two noise terms that are a function of tape magnetization and one noise term independent of the magnetic state of the tape.

Comparison of the theoretical spectrum with measured spectra shows that the noise which increases with magnetization is caused by the distribution of particles, not by interaction between the particles. This comparison also indicates that the spectrum of the noise which increases with magnetization is flat in the frequency range 0-30 kc.

R. F. Bauer

References

1. R. F. Bauer, "A Theoretical and Experimental Study of Noise from Magnetic Recording Tape," Ph. D. Thesis, Department of Electrical Engineering, M. I. T., September 1966.

2. G. J. Fan, "A Study of the Playback Process of a Magnetic Ring Head," IBM J. Res. Develop., Vol. 5, pp. 321-325, October 1961.

3. C. D. Mee, The Physics of Magnetic Recording, John Wiley and Sons, Inc., New York, 1964).

XII. LINGUISTICS[*]

Academic and Research Staff

Prof. R. Jakobson	Prof. C. Watkins	Dr. J. S. Gruber
Prof. J. A. Fodor	Prof. V. J. Zeps	Dr. S.-Y. Kuroda
Prof. M. Halle	Dr. S. Bromberger	Dr. W. K. Percival
Prof. J. J. Katz	Dr. J. V. Canfield	Dr. A. Schwartz
Prof. E. S. Klima	Dr. M. F. Garrett	Dr. D. E. Walker
Prof. G. H. Matthews		J. J. Viertel

Graduate Students

E. W. Browne	J. L. Fidelholtz	P. S. Peters, Jr.
R. J. Carter	M. L. Geis	C. B. Qualls
R. P. G. DeRijk	J. W. Harris	J. R. Ross

A. SURVEY OF LATVIAN MORPHOPHONEMICS

Although the number of alternations encountered in Latvian phonology is rather large, they are determined by a relatively small number of rules.

In this report we shall briefly survey, in the probable order of their application, seven such rules or rule complexes. We shall refer to these as (1) the k/c rule (velars [k, g] are replaced by strident dentals [c, dz] in the environment before front vowels [i, æ]), (2) the i/j rule (diffuse vowels [i, u] are replaced by their corresponding glides [j, w] prevocalically), (3) the n/i rule (a tautosyllabic preconsonantal n is replaced by a diffuse vowel identical in gravity with the preceding vowel), (4) the æ/e rule (æ is narrowed to e before i or j, with or without intervening consonants or glides), (5) metathesis (diphtongs with a compact first vowel [a, æ(e)], metathesize), (6) the V/ϕ rule (a vowel is replaced by zero if the next morpheme starts with a vowel or s, and at the end of a word), and (7) the syncope rule (sequences of identical vowels are replaced by long vowels).

The above rules are applicable to strings assembled in the morphological component; the output of these rules is subject to the application of further rules (mutation of dentals, assimilation, etc.). The more important of these will be discussed at the end of this report.

(1)   The k/c rule.

Velars (k, g) are replaced by strident dentals (c, dz) in the environment before front vowels (i, æ):

$$\underline{k}, \underline{g} > \underline{c}, \underline{dz}/\underline{\quad\quad} \text{ FV}$$

Typical examples of this alternation are pairs like <u>ruoka</u> 'hand' and its diminutive ruociņa, aûgu 'I grow' and the corresponding causative aûdzinu, <u>druska</u> 'bit' and its diminutive drusciņa, and <u>ganu</u> 'I herd' with the related <u>dzæ</u>nu 'I drive'.

Since the <u>k/c</u> rule is very early within the morphophonemics of Latvian, its effects are often overlaid with those of other rules, so that apparent exceptions occur in terminal strings.

First of all, the effects of the subsequent V/ϕ rule (rule 6 in this series), which deletes vowels prevocalically and finally, may obscure the conditions that lead to the <u>k/c</u> alternation: aûdz 'you grow' and aûg 'he grows' (from aûg + æ + <u>i</u> and aûg + <u>a</u>, respectively), <u>ræ</u>dzu 'I see' (from ræ<u>gæ</u> + <u>au</u>), <u>saucu</u> 'I call(ed)' (from <u>saukæ</u> + <u>au</u> [Pres.] or <u>saukæ</u> + <u>a</u> + <u>au</u> [Past]).

Second, in cases where metathesis (rule 5) takes place in addition to vowel truncation (rule 6), it may appear that the environment calls for a <u>k/c</u> alternation, where non occurs: <u>vilki</u> 'wolves' (from <u>vilk</u> + <u>ai</u> > [met.] <u>vilk</u> + <u>ia</u> > [V/ϕ] <u>vilk</u> + <u>i</u>), <u>saki</u> 'you say' (from <u>sak</u> + <u>a</u> + <u>æi</u> > [met.] <u>sak</u> + <u>a</u> + <u>ia</u> > [V/ϕ] <u>sak</u> + <u>i</u>).

Third, the <u>k/c</u> alternation, seemingly implicit in morphological assemblies such as <u>skæl</u> + <u>ti</u> 'to split', may be superseded by the operation of another rule, e. g. , the <u>k/ķ</u> alternation in the environment s <u>____</u> FV (precise statement elsewhere): šķelt.

Fourth, the <u>k/c</u> alternation may be ruled out from some (foreign) lexical items: <u>kino̅</u> 'movies'.

As far as we have been able to ascertain, the <u>k/c</u> rule is not ordered with respect to rules (2), (3), and (4); it must, however, precede rules (5) and (6), as shown by examples aûdz and <u>vilki</u>, above.

(2) The <u>i/j</u> rule.

All native instances of <u>j</u> and <u>v</u> point to their being prevocalic realizations of the vowels <u>i</u> and <u>u</u>, respectively, e. g. , šuvu 'I sewed' and šūt 'to sew', <u>lija</u> 'it rained' and <u>li̅t</u> 'to rain', <u>skreju</u> 'I run' and <u>skriet</u> 'to run'.

Accordingly, no instances of <u>v</u> and <u>j</u> appear in the lexical core. All instances of <u>v</u> and <u>j</u>, at least in the basic morphemes, are due to the glide rule, which states, informally, that

$$\underline{i}, \underline{u} > \underline{j}, \underline{w}/\underline{\quad\quad} \text{ V}$$

or <u>i</u> and <u>u</u> are replaced by their respective glides before vowels (and a late rule then carries <u>w</u> to <u>v</u>).

While the rule must be stated in the entirely general form above, there are a number of environments from which it has been specifically exempted. In particular, the

rule does not apply to any segment that precedes an identical vowel (i. e. , the rule does not apply to the first i in ii or to the first u in uu, and Latvian does not have the sequences ji and vu, except under very special conditions): iilæn + a + s > īlæns 'awl' (not *jilæns) and uudæn + i + s > ūdens 'water' (not *vudens). The rule does not apply to an i before a front vowel in the next morpheme: saaksi + æi > [met.] saaksi + iæ > [V/φ] saaks + i > [sync.] sāksi 'you will begin' (and not *sāksĭ from *sāksj + æi). Nor does the rule apply to an u that follows a morpheme-initial consonant cluster ending in other than a dental obstruent: tvans 'carbon monoxide' but puika 'boy', muita 'customs', kuilis 'boar'.

Further exceptions are either lexical (tuaregs 'Tuareg', diēta 'diet'), or appear in the lexicon with a v and j already specified as non-vocalic: Vladivostoka, jipts colloq. 'poison'.

The i/j rule is not ordered with respect to rules (1), (3), and (4). It must, however, precede metathesis (5) — cf. java 'fluid mixture' and ieva 'bird-cherry' from iau + aa + i and æiu + aa + i, respectively. The i/j rule must likewise precede the vowel truncation rule (6) and the syncope rule (7) — (in order to yield lija 'it rained', the underlying lii + aa may not be replaced by *li + aa or *lī + aa).

(3) The n/i rule.

The alternations covered by this rule involve the conversion of a tautosyllabic pre-consonantal or final n into a high vowel identical in gravity with the preceding vowel. Thus, unC is converted into uuC, ænC into æiC, and so on. The form of the rule is, informally:

$$n > \begin{Bmatrix} i/FV \\ u/BV \end{Bmatrix} \underline{\quad} \begin{Bmatrix} C \\ \# \end{Bmatrix}$$

None of the n/vowel alternations result in sequences that are retained in terminal strings. Some are further subject to the syncope rule: gin + ti > (k/c) dzin + ti > (n/i) dzii + ti > (V/φ) dzii + t > (sync.) dzīt 'to drive', iunt + a > (i/j) junt + a > (n/i) juut + a > (V/φ) juut > (sync.) jūt 'he feels', krint + a > (n/i) kriit + a > (V/φ) kriit > (sync.) krīt 'he falls', lind + aa > (n/i) liid + aa > (V/φ) liid + a > (sync.) līda 'he crawled'.

Other intermediate strings (and these constitute the majority) are transposed by the metathesis rule: zanb + a + s > (n/i) zaub + a + s > (met) zuab + a + s > (V/φ) zuabs > (other rules) zuops 'tooth', lænk + ti > (n/i) læik + ti > (æ/e) leik + ti > (met.) liek + ti > (V/φ) liekt 'to bend', prant + a > (n/i) praut + a > (met.) pruat + a > (V/φ) pruat > (assim.) pruot 'he knows how', brænd + n + a > (n/i) bræid + n + a > (æ/e) breid + n + a > (met.) bried + n + a > (V/φ and other) brien 'he wades'.

The n/i rule (3) has to precede the æ/e rule (4), if the underlying form for Asf zæm + æn 'earth' is correct. The form zæm + æn should first go to zæm + æi, then narrow to

zem + ei, metathesize to zem + ie and truncate to the correct zemi. Since the narrowing is caused by i, it follows that the n/i rule must precede the æ/e rule. [There are other possible explanations for the above; these, however, seem less attractive.] The n/i rule (3) certainly has to precede metathesis (5), as already shown.

Exceptions to the n/i rule are numerous, and all entirely lexical (i. e., there is no subclass of morphemes, foreign words aside, that consistently retain n): rinda 'row', dzintars 'amber', censties 'to strive', tænkas 'gossip', anglis 'Englishman', and many more.

(4)   The æ/e rule.

The narrowing of æ to e takes place before i or j, with or without intervening consonants or glides. Typical are examples like mieta 'maid', leja 'valley', slēpju 'I hide', ceļu 'I lift' (from kæḷi + au). The narrowing, furthermore, is "contagious" insofar as any number of vowels will be narrowed, so long as an a or u doesn't intervene: æcâetu 'I would harrow' and ecêsi 'you will harrow'.

Apparent discrepancies are of two kinds. In one case, it appears as if the conditions have been met, but no narrowing has taken place. The original conditions have been most typically obscured by metathesis and truncation, e. g., dâeli (from dææêel + ai) 'sons'. There are, however, other exceptions, e. g., verbs in -in-: vædinât 'to air'.

Conversely, it may appear that the conditions have not been met, yet the narrowing has occurred. Here we typically have to do with medial or terminal vowel truncation, e. g., mest 'to throw' (from mæt + ti). Except under special conditions (æiẋe, xærmanis 'Eiche, Hermanis'), e in loanwords is narrow: 'leta 'counter', teẋnika 'technology', vetō 'veto'.

As already shown, the æ/e rule must precede metathesis, and follow the n/i rule.

(5)   Metathesis.

In Latvian phonology, metathesis plays a central role. It is unconditioned, i. e., the rule applies to all appropriate sequences regardless of environment, unless specifically blocked. Subject to metathesis are ai, au, au from an, æi, æi from æn, and æu, whereby æi from æn and au from an metathesize invariably.

Exempt from metathesis are: ai in root syllables: maize 'bread', laiva 'boat'; au in root syllables unless before a root-final vowel: saule 'sun' but guovs 'cow' (from gauu + i + s); archaic imperatives eima 'let's go' and eita 'go!'; a small number of highly frequent stems with æi (and their derivatives: beigt 'to end,' beigas 'end', beigts 'dead', teikt 'to say', teika 'legend', teicams 'praiseworthy', etc.; and late loans: streiks 'strike'.

Examples of metathesis include skræi + n + a > skrien 'he runs' vs. skrēja 'he ran', lænk + a > liek 'he puts', dæu + d + a > duod 'he gives' vs. deva 'he gave', vilk + ai >

(met.) vilk + ia > > (V/φ) vilki 'wolves', skaties 'you look' (from skat + a + æisi), lank + a + s > (met.) lauk + a + s > (V/φ and other) luoks 'bow', sit + a + au > (met.) sit + a + ua > (V/φ) situ 'I hit (past)'.

The place of metathesis in the sequence of rules is perhaps best motivated of all. Rules (1) through (4) must precede it (see above), and the V/φ rule must follow it (see below).

(6)   The V/φ rule.

Terminal strings in Latvian may not have numerous successive vowels; to the degree that they have not been converted into glides, excess vowels are truncated. Vowels must be truncated if the next morpheme (from a phonological point of view — i. e., a stretch of segments flanked by pluses) starts with a vowel or s and at the end of the word:

$$V \to \phi \Big/ \underline{\qquad} \begin{cases} + \begin{cases} V \\ s \end{cases} \\ \neq \end{cases}$$

Accordingly, the reflex of an assembly of the type ræg$\bar{æ}$ + au 'I see' is rædz$\bar{æ}$ + au (k/c rule) → rædz$\bar{æ}$ + ua (metathesis) → rædzu (truncation); sāki 'you began', is derived from saak + a + æi (metathesis) → saak + a + iæ (truncation) → saak + i (syncope) → sāki. Truncation before + s is illustrated by vilks 'wolf' from vilk + a + s.

Truncation must follow metathesis, as the examples rædzu and sāki illustrate. It must precede syncope (see below).

(7)   Syncope.

The syncope rule contracts sequences of identical vowels into single long vowels; lii + ti → līt 'to rain'. Its place after the truncation rule is shown by the following example: skat + aa → skat + a 'he views' (truncation), i. e., the syncope rule changing aa to ā (as in skatām 'we view') may not apply prior to the truncation rule.

The rules just cited imply obviously appropriate underlying representations so that the attested phonetic reflexes may be derived. The underlying strings in the major flexional categories of Latvian are of special interest and will be reviewed below. Rules other than the seven just discussed apply as well; they do not, however, affect our conception of the underlying forms.

We illustrate the declension with the assemblies and phonetic realizations of vilk m. a-stem 'wolf', gulb m. i-stem 'swan', lap f. a-stem 'leaf' and zæm f. æ-stem 'earth'.

Nsm	vilk + a + s	vilks	gulb + is	gulbis
Gsm	vilk + a + aa	vilka	gulb + iaa	gulbja
Dsm	vilk + ami	vilkam	gulb + imi	gulbim
Asm	vilk + an	vilku	gulb + in	gulbi
Lsm	vilk + aaia	vilkā	gulb + iiia	gulbī

Npm	vilk + ai	vilki	gulb + iai	gulbji
Gpm	vilk + au	vilku	gulb + iau	gulbju
Dpm	vilk + aimi	vilkiem	gulb + iaimi	gulbjiem
Apm	vilk + au + s	vilkus	gulb + iau + s	gulbjus
Lpm	vilk + ausi	vilkuos	gulb + iausi	gulbjuos
Nsf	lap + aa + i	lapa	zæm + ææ + i	zeme
Gsf	lap + aa + si	lapas	zæm + ææ + si	zemes
Dsf	lap + aia	lapaj	zæm + æia	zemej
Asf	lap + an	lapu	zæm + æn	zemi
Lsf	lap + aaia	lapā	zæm + ææia	zemē
NApf	lap + aa + si	lapas	zæm + ææ + si	zemes
Gpf	lap + au	lapu	zæm + iau	zemju
Dpf	lap + aami	lapām	zæm + ææmi	zemēm
Lpf	lap + aasi	lapās	zæm + ææsi	zemēs

The indefinite adjective is declined like the corresponding masc. and fem. a-stems: Nsm mazs 'small', Gsm maza, etc., Nsf maza, etc.

The definite adjective and the pronoun differ from the noun in one of two ways — either an element ai is infixed (in the Nsm, D, and L forms of the adjective, and in the L forms of the pronoun), or the morpheme containing the thematic a-vowel is exempt from the truncation rule. Paradigms of mazajs 'the small (one)' and tas 'that' follow.

Nsm	maz + aia + s	mazajs	t + a + s	tas
Gsm	maz + a + aa	mazā	t + a + aa	tā
Dsm	maz + aiami	mazajam	t + ami	tam
Asmf	maz + an	mazuo	t + an	tuo
Lsmf	maz + aiaaia	mazajā	t + aiaaia	tajā
Nsf	maz + aa + i	mazā	t + aa + i	tā
Gsf	maz + aa + si	mazās	t + aa + si	tās
Dsf	maz + aiaia	mazajaj	t + aia	taj

Asf as Asm, Lsf as Lsm

Npm	maz + ai	mazie	t + ai	tie
Gpm	maz + au	mazuo	t + au	tuo
Dpm	maz + aiaimi	mazajiem	t + aimi	tiem
Apm	maz + au + si	mazuos	t + au + s	tuos
Lpm	maz + auausi	mazajuos	t + aiausi	tajuos

NApf	maz + aa + si	mazās	t + aa + si	tās
Gpf as Gpm				
Dpf	maz + aiaami	mazajām	t + aami	tām
Lpf	maz + aiaasi	mazajās	t + aiaasi	tajās

In the conjugation, first note the assemblies and realizations of the future tense and its reflexive for the verb <u>likt</u> 'to put':

1s	liksi + au	likšu	liksi + ausi	likšuos
2s	liksi + æi	liksi	liksi + æisi	liksies
1p	liksi + mi	liksim	liksi + mi + æisi	liksimies
2p	liksi + ti	liksit	liksi + ti + æisi	liksities
3	liksi	liks	liksi + æisi	liksies

The reflexive morpheme has two allomorphs. The allomorph + <u>æisi</u> occurs after <u>Ci</u>, the allomorph <u>si</u> (without +) occurs elsewhere. The first person sg. morpheme is <u>au</u>, the second person sg. morpheme is generally <u>æi</u>, but in the present tense non-reflexive assemblies of a number of verbs, the second person sg. is introduced with a morpheme boundary as <u>æ + i</u>. The past tense morpheme is realized as <u>aa</u>, the present tense morpheme as <u>aa</u> or <u>a</u> depending on the stem-clan. Prior to the application of any morphophonemic rules, one <u>a</u> is deleted from the singular assembly (i.e., <u>mæt + a + æ + i</u>, underlying <u>met</u> 'you throw', is adjusted to <u>mæt + æ + i</u>, and <u>saak + aa + æi</u>, underlying <u>sāki</u> 'you began', is adjusted to <u>saak + a + æi</u>). This adjustment is crucial if the <u>k/c</u> rule, the <u>i/j</u> rule, the <u>æ/e</u> rule and the V/ϕ rule are to operate properly. A few paradigms follow, already adjusted as to the extra <u>a</u>. The stems are <u>mæt</u> ϕ-pres. stem 'throw', <u>lænk</u> ϕ-pres. stem 'put', <u>sak ī</u>-past stem 'say'

Present

1s	mæt + au	mætu	lænk + au	lieku	sak + a + au	saku
2s	mæt + æ + i	met	lænk + æ + i	liec	sak + a + æi	saki
1p	mæt + a + mi	mætam	lænk + a + mi	liekam	sak + aa + mi	sakām
2p	mæt + a + ti	mætat	lænk + a + ti	liekat	sak + aa + ti	sakāt
3	mæt + a	mæt	lænk + a	liek	sak + aa	saka

In the past tense, the stem allomorphs <u>met</u>, <u>lik</u>, and <u>sakii</u> are morphologically conditioned:

1s	met + a + au	metu	lik + a + au	liku
2s	met + a + æi	meti	lik + a + æi	liki
1p	met + aa + mi	metām	lik + aa + mi	likām
2p	met + aa + ti	metāt	lik + aa + ti	likāt
3	met + aa	meta	lik + aa	lika

1s	sakīi + a + au	sacīju
2s	sakīi + a + æi	sacīji
1p	sakīi + aa + mi	sacījam
2p	sakīi + aa + ti	sacījat
3	sakīi + aa	sacīja

In addition to the seven rules discussed at the beginning of this report, a number of other rules apply as well. These, however, are all sufficiently close to the phonetic surface, that they present no theoretical or practical difficulties. The most important among these are the following forms

(a)   The t/s alternation.

Dental stops (t, d) are replaced by spirants (s, z) in the environment before other dental stops and before j.

t, d > s, z/_____ t, d, j

e. g. , mæt + ti > (æ/e) met + ti > (V/φ) met + t (t/s) mest 'to throw'. In most cases, the immediate reflex of t and d (i. e. , s, z) is subject to further rules: the mutation of dentals to palatals (b) — Dat. sg. mædi + ami > (i/j) mædj + ami > (æ/e) medj + ami > (V/φ) medj + am > (t/s) mezj + am > (palat.) mežj + am > (j/φ) mežam 'forest'; assimilation in voicing (d) — væd + ti > (æ/e) ved + ti > (V/φ) ved + t > (t/s) vez + t > (assim.) vest 'to lead'; or loss (not discussed in this report) — kliid + tt + a > (V/φ) kliid + tt > (sync.) klīd + tt > (t/s, across a plus) klīz + tt > (t/s) klīz + st > (assim.) klīs + st > (loss[or syncope ?]) klīst 'he strays'.

It is essential that the t/s rule precede the mutation of dentals — cf. zutis 'eel' and Gs zuša (from zut + iaa > (i/j) zut + jaa > (V/φ) zut + ja > (t/s) zus + ja > (palat.) zuš + ja > (j/φ) zuša.

(b)   The mutation of dentals.

Dental consonants are replaced by corresponding palatals in the environment before j:

c, dz, s, z, n, l, r > č, dž, š, ž, ņ, ļ, ŗ/_____ j, e.g. , laak + iaa > (k/c) laac + iaa > (i/j) laac + jaa > (V/φ) laac + ja > (sync.) lāc + ja > (palat.) lāč + ja > (j/φ) lāča 'bear Gen sg.'

The mutation rule may be considerably more general than suggested above and may include the environment "after j" as well (vējš 'wind' from uæææi + a + s) and the environment before other palatals (sķelt 'to split' from skæl + ti). There are, however, some minor difficulties in the formulation of a more general rule, among them the apparent counterexamples, such as maijs 'May' (not *maijš) and runča Gs 'tomcat' (not *ruņča).

(c)   The j/φ alternation.

In a number of environments, a j̠ is lost. These include the environment after long vowels at the end of polysyllabic words, e.g., mazg̅ai̠ + a̠ > (i/j) mazg̅aj̠ + a̠ > (V/φ) mazg̅aj̠ > (j/φ) mazg̅a 'he washes'; as well as the environment after a palatal (the latter generated by the previous rule): naz̠ + iai̠ > (i/j) naz̠ + jai̠ > (met.) naz̠ + jia̠ > (V/φ) naz̠ + ji̠ > (palat.) naž̠ + ji̠ > (j/φ) naži̠ 'knives'.

It is not clear whether the j/φ loss is connected with the loss of other segments, e.g., d̠ before n̠: brænd̠ + n̠ + a̠ > (n/i) bræid̠ + n̠ + a̠ > (æ/e) breid̠ + n̠ + a̠ > (met.) bried̠ + n̠ + a̠ (V/φ) bried̠ + n̠ > (d/φ) brien̠ 'he wades'.

(d)   Assimilation in voicing.

Obstruent clusters are subject to a regressive assimilation in voicing, e.g., zirg + a̠ + s̠ > (V/φ) zirg + s̠ > (assim.) zirks 'horse'. This assimilation must follow j-loss: mædi̠ + a̠ + s̠ > (i/j) mædj̠ + a̠ + s̠ > (æ/e) medj̠ + a̠ + s̠ > (V/φ) medj̠ + s̠ > (t/s) mezj̠ + s̠ > (palat.) mež̠j + š̠ > (j/φ) mež̠š̠ > (voicing) meš̠š̠ 'forest'.

M. Halle, V. J. Zeps

B.   ON THE METRICS OF PRE-ISLAMIC ARABIC POETRY

The purpose of the following note is to give wider currency to certain facts of great linguistic interest which because of the present somewhat artificial organization of scholarly publication are likely to escape the attention of all but a very small number of linguists. The note is based on G. Weil's article ᶜAr̅ud̠ in the Encyclopedia of Islam, I (Leiden, 1960), pp. 667-677.

The term ᶜar̅ud̠ is used by the native Arab grammarians to designate the "science of the rules by means of which one distinguishes correct metres from faulty ones in ancient (pre-Islamic — M. H.) poetry." (667). In line with this conception of the primary objectives of the science of prosody, the ancient grammarians distinguished between us̠̅ul̠, which refers to the abstract underlying patterns, and fur̅uᶜ, which refers to the set of verse types by means of which the abstract patterns may be actualized. If this distinction were to be applied in the study of English prosody, basic verse patterns such as "iambic pentameter" or "trochaic dimeter" would belong to us̠̅ul̠, whereas the rules for actualizing the iambic pentameter — e.g., the list of the "allowable deviations from the basic iambic pentameter" — would belong to the fur̅uᶜ.

In the present note we restrict our attention to the rules that establish the 16 us̅ul̠ patterns recognized by the native prosodists.

Rule 1.   The verse line is composed of two identical hemistichs (miṣr̅aᶜpl. maṣ̠̅ariᶜ).[1]

Rule 2. A hemistich contains either three or four pegs (watid pl. awtad).

Rule 3. In a given hemistich each peg is preceded by the same number of cord units (sabab pl. asbab). This number may vary between one and three.

Rule 4. In hemistichs with four pegs no more than two cord units are admitted before a peg.

Rule 5. In hemistichs with three pegs, no less than two cord units must precede each peg.

If we represent a peg by P and a cord unit by C, we obtain from rules 2-5 the following four abstract patterns, which we designate here by the numerals with which they are labelled in the Arab treatises:

(1)

CPCPCPCP	V
CCPCCPCCPCCP	I
CCPCCPCCP	III
CCCPCCCPCCCP	II

Rule 6. A peg is composed of a weak position and strong position.

In a regular line of verse the weak position is occupied by a short unstressed syllable and the strong position is occupied by a long stressed syllable. When the weak position precedes the strong position we have an iambic peg (watid majmūᶜ); when the strong position precedes the weak position, we have a trochaic peg (watid mafrūk). The occurrence of trochaic pegs is severely restricted (see rule 7 below).

A cord unit is normally occupied by a single syllable whose quantity and stress are apparently free.

Rule 7. In three-peg hemistichs with two intervening cord units (pattern III in (1)) the last peg may be either iambic or trochaic; otherwise trochaic pegs are not admitted, and only iambic pegs are found.

Thus, if we assume that P in (1) stands for an iambic peg, and let Q stand for a trochaic peg, we may think of rule 7 as adding a fifth pattern to the four cited m (1):

(2)   CCPCCPCCQ   (IV)

Rule 8. Subject to the restrictions below (see rules 9-11) the set of admissible hemistichs is given by the strings formed by cyclical permutation from the five patterns in (1) and (2).

Rule 8 expresses the main constitutive principle of Arabic. The founder of Arabic prosody, the eighth century scholar, Al-Xalīl, expressed this fact by representing the five patterns in (1) and (2) as circles. Thus, for instance, pattern V of (1) above was represented by Al-Xalīl in a form that is essentially equivalent to (3)

(3)
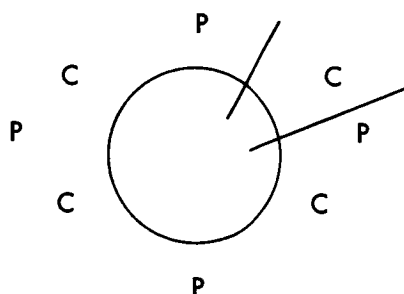
where the two lines intersecting the circle indicate the two possible terminal points of the hemistich. In view of this the five basic patterns of (1) and (2) are referred to traditionally as circles. This term will also be utilized below.

Assuming that the metrical entities are to be read in clockwise order, (3) can be taken as representing the two hemistichs which Al-Xalīl designated as

(4)   mutadārik   CPCPCPCP

         mutakārib   PCPCPCPC

It is obvious that from the pattern of circle V the strings in (4) are the only ones that can be generated by cyclical permutation. From the pattern of circle III in (1) three distinct strings can be generated (as before the names on the left are those used by Al-Xalīl):

(5)   rajaz   CCPCCPCCP

         hazaj   PCCPCCPCC

         ramal   CPCCPCCPC

Similarly, from the pattern of circle I in (1) only three distinct strings can be generated:

(6)   CCPCCPCCPCCP
         PCCPCCPCCPCC
         CPCCPCCPCCPC

These, however, are subject to a special adjustment rule:

Rule 9.   In hemistichs of the first circle delete the cord unit following even numbered pegs, if the line begins with the sequence CP; otherwise, delete the cord unit following odd numbered pegs.

Rule 9 yields then the three attested strings of circle (I):

(7)   basīt   CCPCPCCPCP

         tawīl   PCPCCPCPCC

         madīd   CPCCPCPCCP

Rule 10.   A sequence of three cord units functions in the permutation as a single unit.

Rule 10 accounts for the fact that in circle II of (1) only two strings are generated:

(8)   kāmil   CCCPCCCPCCP

          wāfir   PCCCPCCCPCCC

Rule 11a.   A trochaic peg cannot begin a hemistich.

Rule 11b.   A hemistich may not end in a trochaic peg followed by one or more cord units.

These two rules affect only strings in circle IV.   Rule 11a excludes the string

QCCPCCPCC

Rule 11b excludes the strings

PCCPCCQCC

CPCCPCCQC

The remaining six strings that may be generated by cyclical permutation from circle IV are all admissible:

(9)   sarī^c CCPCCPCCQ

          mujta$\Theta\Theta$ CQCCPCCPC

          muktadab   CCQCCPCCP

          muḍārī^c PCCQCCPCC

          xafīf   CPCCQCCPC

          munsariḥ   CCPCCQCCP

It must be noted that rules 9, 10, 11 have a rather unmotivated appearance in the form in which they are given above.   This suggests that something essential has been missed here.

The preceding discussion differs from the traditional treatment in that it dispenses with the entities foot (jūz^c) and mora, which play a prominent role in the traditional discussions. We have found no use for these entities in our description, and at the present state of our understanding we are inclined to believe that these entities appear in the traditional discussions because of certain features of the symbolic apparatus that is utilized in the traditional discussion, rather than because of properties of the subject matter.

<div align="right">M. Halle</div>

<div align="center">Footnote</div>

1. This term is reminiscent of the Germanic "stave".

C. ON THE INTERSECTION OF REGULAR LANGUAGES AND LANGUAGES GENERATED BY TRANSFORMATIONAL GRAMMARS

1. Introduction

The study[1] of transformational grammars presented in this report is based on the definition of these objects which is forthcoming by Stanley Peters and Robert Ritchie. This definition is in private circulation among workers in the field, and will not be stated here, as it would be twenty pages in length. We shall, however, briefly recapitulate their definition for the present purposes.

A transformational grammar, TG, consists of two parts. The first is a phase-structure grammar which generates bracketed strings of terminal symbols. The second is a sequence of transformations. A transformation is an ordered pair, the first member of which is a structural condition, and the second member of which is a sequence of operations. A structural condition, SC, is a Boolian combination of three predicates on a given factorization of a bracketed string: (i) one sort of predicate says that a certain factor or concatenation of factors is enclosed in a bracket of a certain type; (ii) the second sort says that the interior of a certain factor in the factorization is identical with the interior of another factor; and (iii) the third sort says that the debracketization of a certain factor is identical to a given terminal string. The operations performed by a transformation are of three elementary kinds: (i) the operation of deleting a certain factor; (ii) the operation of replacing a certain factor by a sequence of other factors; and (iii) the operation of adjoining a sequence of factors to a certain factor. The sequence of transformations in the transformational component of the grammar apply by convention to the innermost S-bracket of a string 'working out' to the outermost.

The principle of recoverability of deletions is a constraint on the operations that transformations in the transformational component of the grammar may perform. The principle requires, in essence, that if the ith factor is deleted or is substituted for by the $j^{th} - k^{th}$ factors, then, either the ith factor must be identical to some other factor which is not itself deleted or substituted for by the transformation, or else it must be identical to one of a fixed finite set of terminal strings. It was thought (see e.g., Katz and Postal, An Integrated Theory of Linguistic Descriptions) that placing this constraint on CF-based TG's would make any language generated by such a device fully recursive. Peters and Ritchie have shown that context-sensitive based TG's which meet the condition on recoverability of deletions generate a class of languages which (modulo Church's thesis) is coextensive with the class of all recursively enumerable languages.

We shall be concerned with the implications consequent upon a positive solution to the intersection problem (IP) for transformational grammars (TG). The IP is as follows: Given a context-free based TG, G_1, which meets the condition on recoverability of deletions, and a regular language, H, does there exist a TG, G_2, which is CF-based and

meets the condition on recoverability, such that $L(G_2) = L(G_1) \cap H$?

Note first that a negative solution to this problem implies that the intersection of a CF-based recoverable TG with a CF, context-sensitive or TG language is not generated by another CF-based recoverable TG. As we shall see, however, the interest of IP rests only peripherally on the implications of its negative solution. In essence, a positive solution to IP means that there are several linguistically important properties of CF-based recoverable TG's that hold.

We shall consider TG's that meet the condition on recoverability of deletions and have a CF base only.

2. The Filter Problem

Chomsky[2] states two functions performed by transformations in the construction of grammars for natural languages: the first is "to convert an abstract deep structure that expresses the content of a sentence into a fairly concrete surface structure that indicates its form."[3]; and the second is to act as filters on strings generated by the context-free base allowing only a subset of those strings to qualify as deep structures for sentences grammatical in the language.[4] The formal device that he suggests for accomplishing this is that if some obligatory transformation which should have applied did not because its structural condition was not fulfilled, then a certain symbol called 'sentence boundary' is left in the interior of the string. By convention, then, we accept no strings in which sentence boundaries lie in their interiors.

It is evident that we can state the result of this convention formally by intersecting the transformational language under consideration with a regular language, H, which consists of all those strings over the terminal vocabulary which do not have sentence boundaries occurring in their interiors.

Let us formalize the discussion above.

Definition 1: A filter is a transformation that must apply in each transformational cycle in the derivation of a string in order for that string to be considered an element of the language under consideration. (Notice that this is not a precise statement of what we need linguistically, for, in fact, transformations act as filters in natural languages usually when a certain part of their structural condition is fulfilled, but another part, e.g., an identity condition, is not. It is immediately evident, however, that we can accomplish this result by breaking the transformation in question down into two transformations which perform the same operation and such that one of them is a filter by the above definition.)

Definition 2: The filter problem (FP) is as follows: Is the class of languages generated by TG's with filters larger than the class of TG languages?

LEMMA 1: A positive solution to IP implies a negative solution to FP. (IP $\Longrightarrow$ -FP)

Proof: Let G_1 be a TG, a subsequence of the transformations of which, $T_{i_1} \cdots T_{i_k}$, are filters. Form a TG, G_2, without filters as follows: the base components are the same

except that new symbols $\beta_{i_1} \ldots \beta_{i_k}$ are generated on the right-hand of each S-phrase (i. e. , if $S \rightarrow \omega_j$ is a rule of P_1, then $S \rightarrow \omega_j \beta_{i_1} \ldots \beta_{i_k}$ is a rule of P_2). The transformations are the same, except that for each filter T_{i_j} we add onto its SC the condition that it senses the symbol β_{i_j}, and to its operation an erasure of β_{i_j}. Let H be the regular language over the enlarged terminal vocabulary consisting of those strings in which none of the symbols $\beta_{i_1} \ldots \beta_{i_j}$ ever occurs. It is evident that $L(G_1) = L(G_2) \cap H$.

Notice that by the lemma, then, a positive solution to IP is the 'happy' solution linguistically, in the sense that the linguist is not allowing an increase in the class of structures which he is considering as potential grammars by the use of filters. On the other hand, even a constructive solution to IP, i. e. , one such that we give a procedure for constructing G_2, given G_1, and H, would not mean that we could eliminate filters from use in the present framework of a transformational study of language, for the strong generative capacity would evidently differ between G_1 and G_2. And it is linguistically important that the terminal strings have the correct surface structure, e. g. , for input to the stress assignment rules of the phonological component.

3. The Optional-Obligatory Problem (OOP)

In the definition of TG by Peters and Ritchie, a transformation must apply if its SC is satisfied by some analysis of the string. We have shown above that we can restrict our language to those strings in which a certain transformation, in fact applied in each cycle, modulo the IP. A natural question then concerns the class of languages generated by TG's in which one or more transformation applies optionally. We formalize this notion and the OOP as follows.

Definition 3: A transformation is <u>optional</u> when its application implies its SC has been met. (It is <u>obligatory</u> when it applies iff its SC has been met.)

Definition 4: The <u>optional-obligatory problem</u> is as follows: Does the class of languages generated by TG's in which one or more transformation is optional properly include the class of languages generated by TG's, all of whose transformations are obligatory?

LEMMA 2: A positive solution to IP implies a negative solution to OOP. (IP $\Longrightarrow$ -OOP.)

Proof: We shall consider a TG, G_1, in which only one transformation is optional, and it will be evident that the method of proof can be generalized to the case in which an arbitrary subsequence of transformations of G_1, including all of them, is optional.

Let $G_1 = (P, (T_1 \ldots T_k))$, where T_i is optional. Let $G_2 = (P', (T_1 \ldots T'_i \ldots T_k))$, where P' is like P, except that for every rule in P of the form $S \rightarrow \omega_j$ introduce in addition a rule in P' $S \rightarrow \omega_j \beta_i$; and let T'_i be like T_i, except that in the SC of T'_i we add the condition that the symbol β_i appear on the right and be erased when T'_i applies. Let H be all strings over the enlarged terminal vocabulary in which β_i never appears. We

claim that $L(G_1) = L(G_2) \cap H$. This can be seen by checking cases.

<u>Case I</u>: T_i applied in a certain cycle in G_1. In this case, the symbol β_i was generated by P' and T_i' applied in G_2.

<u>Case II</u>: The SC of T_i was satisfied in a certain cycle, but T_i by option failed to apply. Corresponding to this in G_2, all the SC of T_i' was satisfied, except that the symbol β_i failed to appear as it did not happen to be generated in that cycle. Thus, T_i' failed to apply as its SC was not completely fulfilled.

<u>Case III</u>: The SC of T_i was not fulfilled, and therefore it did not apply in G_1. Then, in G_2 either β_i was generated by P' or it was not. In the latter case, nothing is altered. In the former case, the symbol β_i will not be erased by T_i' and will therefore appear as a terminal symbol; however, this string will not appear in the intersection of $L(G_2)$ and H, since it does not lie in H.

In fact, it would appear that we can arrive at a negative solution to OOP directly without the necessity of a positive solution to IP. In the proof above, add a transformation T_{k+1} which applies in the environment $X\beta_i Y$, and which deletes β_i. Then β_i will never be left unerased in any string generated by G_2, and it is evident that $L(G_1) = L(G_2)$. Thus we can always transfer optionality in the transformations into optionality in the base.

4. The Outermost S-Bracket Problem

According to the definition of Peters and Ritchie, it is impossible for the transformations to 'know' in which S-cycle they are operating in a string generated from the base; in particular, they cannot 'know' when they are in the final S-cycle. From a linguistic point of view this is a deficiency, for there is good evidence that in addition to the cyclical transformations, i.e., the transformations that apply on each S-level from the innermost one out, there must also be a set of postcyclic transforms that apply only at the outermost S-bracket. Furthermore, these postcyclic transformations are interspersed in order among the cyclic transformations in application at the outermost S-bracket.

<u>Definition 5</u>: A <u>postcyclic transformational grammar</u> is a transformational grammar in which a subsequence, $T_{i_1} \ldots T_{i_k}$, of the transformations applies only at the outermost S-bracket if they apply at all to a string.

<u>Definition 6</u>: The <u>outermost S-bracket problem</u> is as follows: Is the class of postcyclic transformational languages properly larger than the class of transformational languages?

<u>LEMMA 3</u>: Let G_1 be a postcyclic TG, then a positive solution to IP implies that there exists a TG, G_2, such that $y \in L(G_1)$ iff $yt \in L(G_2)$, where t is a single symbol not in V_T.

<u>Proof</u>: Let $T_{i_1} \ldots T_{i_k}$ be the postcyclic transformations of G_1. Form G_2 by putting an

additional rule $S \rightarrow \omega_i t$ in P_2 whenever the rule $S \rightarrow \omega_i$ was in P_1. Add the postcyclic transformations to the transformational component of G_2, preserving the relative order of all transformations as they applied on the last cycle in G_1, except that the structural condition of the postcyclic transformations are altered so that they will apply only in the environment Xt in addition to the rest of their SC's. Now, let H be the regular language consisting of all strings over the enlarged terminal vocabulary in which the symbol 't' occurs at the end and only at the end of each string.

The claim is that $y \in L(G_1)$ iff $yt \in L(G_2) \cap H$. For let $y \in L(G_1)$, where y was derived transformationally from the string y' which was generated by the base P_1. We then generate (y')t in the base P_2, and exactly the same transformations which applied to y' will apply to (y')t, including the postcyclic transformations that will apply in the outermost S-bracket of (y')t because the presence of 't' triggers their application in G_2 as we have defined it above. Also, $yt \in H$ by the definition of H. Now let $x \in L(G_2) \cap H$. Since $x \in H$, we know x = yt, where y contains no occurrences of the symbol 't'. Let yt be transformationally derived from the base structure (y')t. We know that $y' \in L(P_1)$ by the construction of P_2. And by the definition of the transformational component of G_2, we know that the same sequence of transformations will apply to (y')t in G_2 as will apply to y' in G_1.

This lemma can be strengthened by using the same methods of proof. That is, all that we have required of the postcyclic transformations in that they apply if they can, but only to the outermost S-bracket. Suppose we require, however, that in addition all the postcyclic transformations must apply in the last cycle (i.e., a string will be in the language iff all of the postcyclic transformations did in fact apply in the last cycle). We might call these postcyclic filters, i.e., transformations which say that the final form of the string must meet such-and-such structural conditions and must have had such-and-such operations applied to it in the last cycle.

The question, as usual, is whether postcyclic filter grammars are essentially stronger than grammars defined by Peters and Ritchie. And the answer, as usual, is that a positive solution to IP implies a negative solution to this problem. The proof utilizes the same methods as the proofs of Lemmas 1 and 3, and may be left as an exercise.

It may be of interest to note that by using intersection with a regular language we could demand that certain transformations apply, e.g., to all S-brackets nested, say, 17 levels below the outermost S-brackets; or that we can demand that transformation T_i apply exactly 5 cycles later than transformation T_j; etc. We shall not prove these assertions, although the proofs are not difficult; rather, the point of the remark is to indicate the extreme power afforded us by intersecting a transformational language with a regular language. Note, further, that the regular languages used are of a rather simple kind. The word 'simple' here can be given a precise meaning in terms of the

number of nestings of cycles generated by a regular grammar. In this sense, the grammars for the regular languages that we have been using are the simplest of all regular languages.

We turn now to the main result of this paper. There is good evidence, as indicated above, that we need to intersect the strings generated by a transformational model of the grammar of a natural language with a regular language in order to produce exactly the grammatical sentences of the natural language. We should therefore want to find a characterization of the class of languages generated by the intersection of a transformational and a regular language. The answer is that we can generate essentially any recursively enumerable set in such a manner.

THEOREM 1: There exists a fixed regular language H such that for any recursively enumerable set of natural numbers S there exists a transformational grammar G such that $\underline{n} \in S$ iff $a^n t \in L(G) \cap H$.[5]

The proof relies on a lemma that is due first to Leonard Haines.

LEMMA 4 (Haines): There exists a fixed context-free language L_1 and a fixed homomorphism λ such that for any recursively enumerable set of numbers S there exists a context-free language L_2 such that $\lambda(L_1 \cap L_2) = \{a^n / n \in S\}$.

The proof of this lemma is forthcoming by Haines, and we shall not repeat it here. We must describe, however, the construction of L_1 and λ for the proof of the theorem.

Let K be the language over the terminal alphabet $\{c, \ldots g\}$ consisting of all those strings which reduce to ϵ (if x reduces to ϵ we shall write R(x)). A string x over this alphabet reduces to ϵ iff it meets one of the following conditions:

(i) $x = \epsilon$

(ii) $(\exists i)(x = cd^i c \, y \, ce^i c)$ & R(y)

(iii) $(\exists i)(x = cf^i c \, y \, cg^i c)$ & R(y)

(iv) $(x = y \, z)$ & R(y) & R(z).

It is clear that we can generate K by some context-free grammar.

Then L_1 is the language over the terminal alphabet $\{a, b, c, \ldots g\}$ as follows: $L_1 = (K \cup \{a, b\}*)*$.

λ is defined as follows: $\lambda(a) = a$;

$$\lambda(b, \ldots g) = e, \text{ where } \underline{e} \text{ is the null string.}$$

L_2 is some CF language over the alphabet $\{a, \ldots g\}$.

The outline of the proof is as follows: In the base P of the grammar G we generate two strings side by side, one each from L_1 and L_2, and in addition an adjoining string from the alphabet $\{a, \ldots g\}$. Let us call these strings x, y, z. The transformations then essentially check, one symbol at a time, that $x = y = z$, and progressively delete x and y at the same time. An additional transformation in each cycle performs the operation of λ on the string z. There is a postcyclic filter that ensures that z is not merely a

proper initial part of the reflections of x and y. If some transformation fails to apply, the string is filtered out. All deletions are in accordance with the principle of recoverability, since we delete items from x and y only as they are equal to something in z, and we delete symbols in z in performing λ by virtue of equality with one of a finite number of terminal strings.[6] A trick is used to ensure that P provides us with the number of S recursions needed to completely refine any string by means of the transformations.

We shall proceed with the details of the description of G. $G = (P, (T_1, T_2, T_3))$, where P is the context-free base defined by the following rules.

$$S \longrightarrow Sa\beta_1 \qquad S \longrightarrow Sa\beta_1\beta_3t \qquad S \longrightarrow S_1S_2a\beta_1$$
$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$
$$S \longrightarrow Sg\beta_1 \qquad S \longrightarrow Sg\beta_1\beta_3t \qquad S \longrightarrow S_1S_2g\beta_1$$

$$S \longrightarrow S_1S_2a\beta_1\beta_3t$$
$$\vdots$$
$$S \longrightarrow S_1S_2g\beta_1\beta_3t$$

This system of 28 rules can be collapsed into one rule by the well-known convention as follows:

$$S \longrightarrow \left\{ \begin{array}{c} S \left\{ \begin{array}{c} a \\ \vdots \\ g \end{array} \right\} \beta_1 \ (\beta_3 t) \\[2em] S_1 S_2 \left\{ \begin{array}{c} a \\ \vdots \\ g \end{array} \right\} \beta_1 \ (\beta_3 t) \end{array} \right\}$$

S_1 is the initial symbol of the CF grammar that generates language L_1 described above, and S_2 is the initial symbol of the CF grammar that generates language L_2, where L_2 varies according to the recursively enumerable set S.

We shall not write the transformations in the formal notation of Peters and Ritchie, since this notation is somewhat opaque. Rather, we shall use an informal notation close to that which appears in most monographs in linguistics. It can be checked that the operations described informally below do not exceed the capacity of the Peters-Ritchie transformations.

T_1.

Analysis: $[\ x\ ]_{S_1} \ [\ y\ ]_{S_2} \ z\ \beta_3\ t$

such that (i) x and y are single terminal symbols

(ii) Z is any string over the alphabet.

Operation: delete 'β_3'.

T_2.

Analysis: $[\ X\ x\ ]_{S_1}\ [\ Y\ y\ ]_{S_2}\ Z\ z\ \beta_1\ W$

such that   (i) X, Y, Z, and W are any strings, including the null string

           (ii) x, y, and z are single terminal symbols

           (iii) x = y = z.

Operation: delete x, y, and 'β_1'.

T_3.

Analysis: Z z (t)

such that   (i) Z is any string, including the null string

           (ii) z is a single terminal symbol from the set $\{b, \ldots g\}$

Operation: delete z.

(Note that T_3 is written as representing really two transformations, one of which operates in the environment 'Z z t', and the other of which operates in the environment 'Z z'.)

Explanation: The language generated by P consists of strings of the form $[_S \cdots [_S [_{S_1} x]_{S_1} [_{S_2} y]_{S_2} u_1 \beta_1 (\beta_3 t)]_S \cdots u_n \beta_1 (\beta_3 t)]_S$, where $x \in L_1$, $y \in L_2$, and $(u_1 \ldots u_n)$ is an arbitrary string over the terminal alphabet. T_1 operates as a postcyclic filter because of construction of the regular language H below. T_2 checks that the rightmost symbol of 'x' equals the rightmost symbol of 'y' equals $\underline{u}_i$ in each cycle. If this equality holds, the former two and the symbol 'β_1' is deleted. If the equality fails to hold, 'β_1' is left in the interior of the string and it is filtered out by construction of H. T_3 performs the operation of the homomorphism λ in each cycle; it is important that T_3 follow T_2.

Let H consist of all strings over the terminal alphabet of the form $a^n t$, all $\underline{n}$. In an accepted string, '$\beta_3 t$' is generated nowhere in the string except in the first cycle. If it appears in the middle somewhere, or not at all, the string is filtered by H. The string is accepted if $x = y = (u_n \ldots u_1)$. T_2 progressively deletes 'x' and 'y', and T_3 performs $\lambda(u_1 \ldots u_n)$. The following cases exhaustively treat strings generated by P which fail to be in $L(G) \cap H$.

Case 1. $lh(x) < n$, or $lh(y) < n$. Then, is some cycle the SC of T_2 will not be fulfilled, T_2 will fail to apply, 'β_1' will not be erased, and the resulting string will not be in H.

Case 2. $lh(x) = lh(y) = n$, but the three-termed equality fails at some cycle. Then T_2 fails to apply, 'β_1' is not erased, and the resulting string is not in H.

Case 3. $lh(x) > n$ or $lh(y) > n$. Then T_1 will sense on the last cycle that there is more than one symbol left in either 'x' or 'y', it will fail to erase 'β_3', and the resulting string will not be in H.

We have, thus, that for all x, $x \in \lambda(L_1 \cap L_2)$ iff $x \in L(G) \cap H$. (For ease of reading, 'β_1' should be written 'β_2', and 'β_3' written 'β_1'. The incongruity is due to a

modification in ordering of the transformations while the proof was being written.)

Consider now the class of CF-based TG's with arbitrary deletions. A member of this class can be considered as a finite calculating procedure, and thus, by Church's thesis, the class of languages generated by this class of grammars is no greater than the set of all recursively enumerable languages. From these considerations, we then have the following result.

COROLLARY 1: For every CF-based TG with arbitrary deletions, G_1, there exists a regular language H and a CF-based TG which meets the condition on recoverability of deletions such that $L(G_1) = L(G_2) \cap H$.

The answer to the converse question is given in the following theorem.

THEOREM 2: For every recursively enumerable set of natural numbers, S, there exists a CF-based TG with arbitrary deletions such that for all $\underline{n}$, $n \in S$ iff $a^n t \in L(G)$.

Proof: The proof proceeds by imitating the proof of Theorem 1 with a slight modification. That is, we add to the grammar described in the proof above a new transformation, T_4, which is ordered to follow the other transformations. T_4 will delete all terms of its analysis if it ever senses 'β_1' or 'β_3', or if it senses 't' in the interior of a string. Actually, as transformations are now set up, it may be impossible to tell a transformation, "do such-and-such if you sense a certain symbol anywhere besides on the rightmost side of the S-phase in which you are working." If 't' is generated by the base of G somewhere in the interior of a string, however, and 'β_3' is in fact erased by T_1 because the latter's SC is satisfied, then on the next cycle 't' will appear either two or else four symbols in from the right-hand side of the S-phase, because of the definition of P. This predictability in the occurrence of 't' allows us to build a mechanism into T_4 to sense 't' if it ever occurs in the interior of a string.

COROLLARY 2: For every CF-based TG_1 which meets the condition on recoverability and every regular language H there exists a CF-based TG_2 over the same alphabet with arbitrary deletions such that $L(G_1) \cap H = L(G_2)$.

The implications of Corollaries 1 and 2 for linguistic theory are as follows: If we increase the power of a transformational model of a grammar of a natural language by allowing intersection with regular languages, as evidently we must in order to handle filters and postcyclic transformations, then we do not reduce the weak generative capacity of the class of grammars proposed as models by restricting them with the condition of recoverability of deletions. The important side of this equivalence is given by Corollary 1; namely, given a language L generated by a CF-based TG with filters and postcyclic transformations and which uses arbitrary deletions, then we can generate L with a CF-based TG with filters and postcyclic transformations which meets the condition on recoverability of deletions.

In terms of weak generative capacity, then, adding the principle of recoverability of deletions to a CF-based TG with filters and postcyclic transformations is ornamental.

It would be incorrect, however, to conclude that we have therefore a mathematical proof that in the present context of transformational analysis the principle of recoverability of deletions is without empirical content, since the linguist is interested in the strong generative capacity of those grammars that he proposes as models for the grammars of natural languages.

Let us return finally to that with which we began, i. e., IP. Corollary 2 suggests a method by which to prove the negative of IP by proving that CF-based TG's with arbitrary deletions are stronger than CF-based TG's that meet the principle of recoverability of deletions. On this basis, one suspects immediately that IP does, in fact, receive a negative solution, although I have not seen a proof that the condition on recoverability of deletions may not be dispensed with where filters and postcyclic transformations are not employed.

Note, however, that a negative solution to IP does not mitigate the force of the remarks above concerning the power that we are adding to a TG by allowing filters and postcyclic transformations. For Theorem 1 could be rephrased by dropping the part about the regular language H and adding that G is a TG with filters and postcyclic transformations. That is, if the conjecture above does in fact hold, namely, if CF-based TG's with arbitrary deletions are stronger than CF-based TG's with recoverable deletions, then we are increasing the generative power of this latter class of objects by adding filters and postcyclic transformations to them.

We should make the final remark that if the IP receives a negative solution, as we have conjectured that it does, then languages generated by the Peters-Ritchie transformational grammars turn out to be strange mathematical objects. For it is known that the other classes of grammars that have received extensive study, namely, regular grammars, context-free grammars, and context-sensitive grammars, do not increase in generative capacity upon intersection with a regular language.

If the IP does, in fact, receive a negative solution, then we have a characterization of the extent to which the addition of the condition of recoverability of deletions restricts the weak generative capacity of CF-based TG's. This is to say, dropping the condition on recoverability will extend the weak generative capacity of CF-based TG's exactly as much as will intersecting the 'recoverable' languages with regular languages. Furthermore, if IP receives a negative solution, then allowing filters and postcyclic transformation extends the generative capacity of CF-based TG's exactly as much as does intersecting the languages generated by these devices with regular languages. A stronger result, which comes from an analysis of the proof of Theorem 1, is that the addition of filters and postcyclic transformation extends the generative capacity of CF-based TG's exactly as much as the generative capacity of these grammars is restricted by adding the condition on recoverability of deletions, providing that IP has a negative solution.

5. Linguistic Implications

It may be plausibly argued that the results obtained concerning the generative power of certain classes of transformational grammars are of only peripheral interest for linguistics. For, what is of interest linguistically is the internal structure of the recursive devices that serve as models for grammars. That is, suppose that we find (somehow) that we do need a class of devices to serve as models of grammars such that any recursively enumerable set can be generated by one of these devices. This does not mean that the problem of constructing grammatical models for natural languages loses interest on the ground that all we shall have shown is that grammars are, after all, finite computational devices, which is always the implicit assumption in any case. The linguistically interesting properties of grammars reside largely in the formal properties of their rule systems, rather than in their generative capacity.

Generative capacity has been of linguistic interest, in that past, only in a negative sense. That is to say, a characterization of the generative capacity of a class of devices has been useful traditionally only in showing that such kinds of devices can not function as models for grammars. Thus, it can be shown that there are nested dependencies obtaining in natural languages which in principle are beyond the expressive capacity of regular grammars, and presumably the same is true of grammars that generate context-free languages, although several proofs of this now in circulation need revision, since Ullian has shown a conjecture of Haines concerning nondoubling languages to be false.

In this light, we may ask whether the results reported here are of any central interest to linguistics. I believe they are in the sense of suggesting certain changes in the theoretical structure of Aspects. The proof of Theorem 1 does not depend upon deep properties concerning the internal structure of TG's, but rather relies on a series of tricks. These tricks are predicated on the fact that a transformation can read a certain terminal symbol generated by the base which triggers its operation. Note also, that we used this device in the proof of Lemma 2, which is linguistically surprising in view of the past discussions concerning whether transformations in natural language ought to be obligatory, optional or a mixture. The current work in restricting the generative capacity of TG's is concerned with strengthening the condition on recoverability of deletions.[7] One very obvious means, however, of restricting generative capacity which is suggested by the theorems above is that of not allowing transformations to be sensitive to the existence of terminal symbols or terminal strings. Formally, this would mean that we eliminate one kind of predicate from availability for use in the SC of a transformation, namely, the third kind mentioned above.

The constraint of not allowing transformations to 'read' lexical strings has been proposed by others. And we find in our results further reasons of a purely mathematical nature for imposing this constraint. The empirical question that is raised, then, is that

of whether a transformational grammar with this constraint is sufficiently powerful to generate all and only the grammatical sentences of a given natural language.

J. P. Kimball

Footnotes and References

1. I am indebted to Dr. Leonard Haines for many discussions on the material which is presented here. In particular, he suggested the use of his theorem, which appears as a lemma, as a tool by means of which to prove the main result of this report. I am, of course, solely responsible for any errors.

2. N. Chomsky, Aspects of the Theory of Syntax (The M. I. T. Press, Cambridge, Mass., 1965).

3. Ibid., p. 136.

4. Ibid., pp. 138-139.

5. A result obtained subsequent to the writing of this report strengthens this result to the extent of showing that the TG in question may have a regular grammar as a base instead of a context-free grammar. That is, the class of CF-based TG's is no stronger than the class of regular-based TG's. The proof is too lengthy to be included here, but will be submitted for publication elsewhere.

6. Recent discussion concerning syntax has suggested that the condition on recoverability of deletions should be strengthened to require that if terminal string $'t_1'$ is deleted by virtue of identity to terminal string $'t_2'$, the identity must be one of deep structure in addition to simple identity of the debracketed strings. We should then want to know if strengthening the condition in this way restricts weak generative capacity. Let us define the deep structure of a terminal string $'t'$ informally to be the lowest node which dominates all of $'t'$ in the string generated by the base. Following this definition, it is easy to show that the proof of our main theorem can be carried through under this stronger condition on recoverability of deletions. We change the base component so that if a terminal symbol is generated by the rule $A \rightarrow X b Y$ in the base, then instead it will be generated by the two rules $A \rightarrow X B Y$, $B \rightarrow b$, where B is a new nonterminal symbol. Thus, by definition, every occurrence of the terminal $'b'$ will have the same deep structure, to wit, $'B'$. We can see that the proof of the main result above goes through under the stronger condition on recoverability of deletions by noting that only terminal strings of length one are deleted by any single application of a transformation.

7. A. B. Peters, personal communication, 1966.

XIII. COGNITIVE INFORMATION PROCESSING*

Academic and Research Staff

Prof. S. J. Mason
Prof. W. L. Black
Prof. M. Eden
Prof. T. S. Huang
Prof. F. F. Lee

Prof. W. F. Schreiber
Prof. O. J. Tretiak
Prof. D. E. Troxel
Dr. M. P. Beddoes
Dr. D. Cohen

Dr. P. A. Kolers
Dr. N. Sezaki
C. L. Fontaine
K. R. Ingham
G. L. Wickelgren

Graduate Students

G. B. Anderson
A. K. Bhushan
P. K. Bice
J. D. Bigham, Jr.
D. Caldwell
A. L. Citron
R. W. Cornew
A. Gabrielian

R. V. Harris III
H. P. Hartmann
R. W. Kinsley, Jr.
M. B. Lazarus
J-H. Liu
J. I. Makhoul
H. A. Mildonian

J. A. Newell
L. C. Ng
C. L. Seitz
S. D. Shoap
A. Spiridon
R. M. Strong
R. M. Tate
J. A. Williams

A. PICTURE PROCESSING

1. SCANNER DISPLAY (SCAD)

Purpose of the Equipment

The Scanner-Display system, called SCAD, is a device for communicating between computers and pictures. In its present form it can accept transparencies, measure the transmission at various points on the transparency, and convert this to a digital number. It can also accept digital data and produce pictures from it.

The device is a transducer between digital computers and pictures. SCAD can be operated with digital tape or with a computer. When SCAD is operated with tape, it can either scan transparencies and write the digital information on a magnetic tape, or it can read a digital magnetic tape and display the data from it on a CRT. The display can be viewed and photographed. When SCAD is operated by a digital computer, the computer can either interrogate the transmittance of a transparency at points chosen by the computer, or it can display information, again with brightness and coordinates chosen by computer. The scanner can measure color, as well as brightness.

SCAD measures brightness and displays data over a raster of discrete points. The raster may contain 1024 points along each of two orthogonal dimensions, although fewer points than this may be used. The resolution of the scanner and display depends on the components used in these units, and may be more or less fine than the spacing of the

raster points. Brightness information can be quantized to 8 bits (256 levels).

Block Diagram

A symbolic block diagram of SCAD is shown in Fig. XIII-1. The transparency to be scanned is placed in the scanner. The picture being displayed by the system may be viewed on the monitor or photographed with the display unit. The deflection amplifiers accept the deflection signal from the central control unit, and generate deflection currents for the scanner, display, and monitor. The analog section generates voltages for controlling the intensification of the scanner, the display, and the monitor, and accepts and provides control and brightness signals from and to the central control unit.
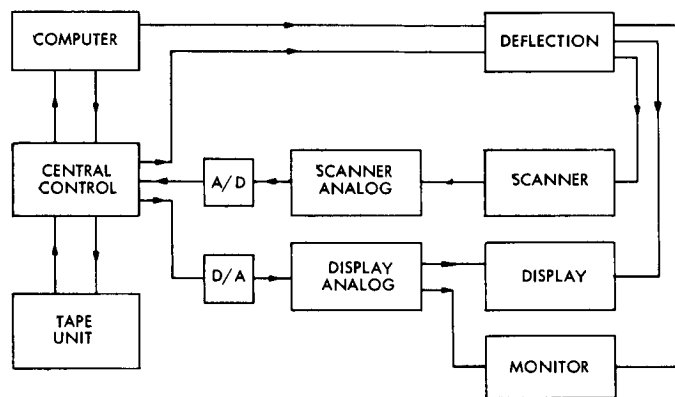
Fig. XIII-1. SCAD system diagram.

The central control unit is the clearing house for data flow between the analog section and the tape unit or the computer. It also sequences the operation of the deflection in the scanner and the display. When SCAD is being operated with a tape unit, central control also generates signals that control tape motion.

a. Scanner Block Diagram

A symbolic diagram of the Scanner is given in Fig. XIII-2. A spot of light from the cathode-ray tube face is imaged on a transparency. The light passing through the transparency is collected by the condenser and is sent through a sequence of dichroic mirrors. The mirrors split the light spectrum into three bands, and each band is sent to a different photomultiplier.

To counteract variations in the light given off by the cathode-ray tube part of the light coming from it is taken by a beam splitter and sent to a photomultiplier tube. The signal from this photomultiplier is used to control the CRT drive.
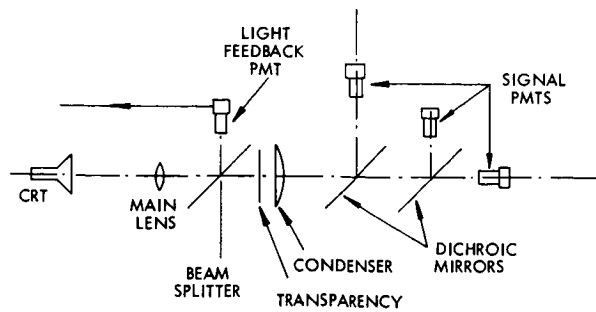
Fig. XIII-2. Scanner symbolic diagram.

b. Display Unit

A symbolic diagram of the Display Unit is shown in Fig. XIII-3. The Display cathode-ray tube is mounted underneath a Polaroid MP-3 copy camera. A partially reflecting mirror is interposed between the camera and the CRT and a portion of the light coming from the CRT falls on a photomultiplier tube. The signal from the photomultiplier tube may be used to adjust the drive to the CRT.
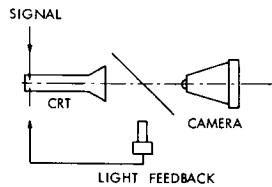

Fig. XIII-3. Display unit.

c. Monitor

The Monitor consists of just a cathode-ray tube with a long-persistence phosphor. The same brightness information that is supplied to the Display is also displayed on the Monitor.

d. Analog Section

The analog section associated with Scanner operation amplifies the signal from the scanner, puts this through an appropriate nonlinear amplifier, and presents this signal to the A/D converter. It also generates the voltages necessary to operate the Scanner cathode-ray tube.

The Scanner Amplifier portion of the analog section is shown in Fig. XIII-4. The scanner is operated in intermittent fashion: the CRT is normally blanked, and it is
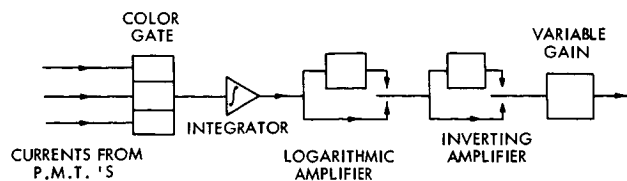

Fig. XIII-4. Scanner analog amplifiers.

unblanked only when one wishes to measure the transmittance of some point on the transparency. The anodes from the three pickup PMT come into a gate which permits current from only one PMT to enter the integrator. The tube is blanked at the end of the intensification period. The output of the integrator is fed into an amplifier chain that can provide one of four transfer characteristics. Let S be the final signal output, T the transmittance of the transparency, and k a variable, proportional to the gain of the variable gain amplifier, and A a constant equal to the peak value of the signal. The four transfer functions are

1. Linear normal

$$S = kT$$

2. Linear inverted

$$S = k(A-T)$$

3. Log normal

$$S = k(\log T)$$

4. Log inverted

$$S = k(A - \log T)$$

Log normal is the usual transfer function for operating the scanner. The "contrast" is controlled with the variable amplifier gain. The "brightness" is adjusted by setting the PMT supply voltage.
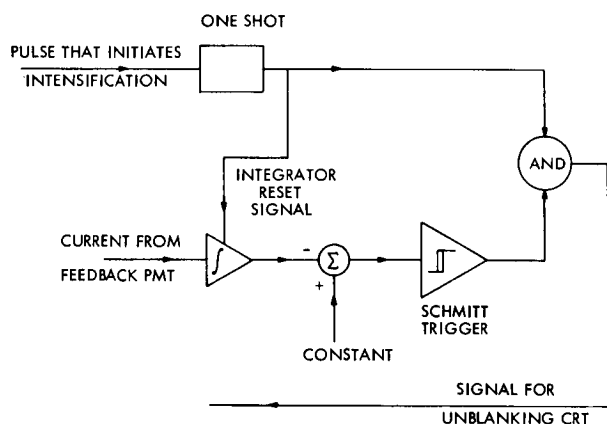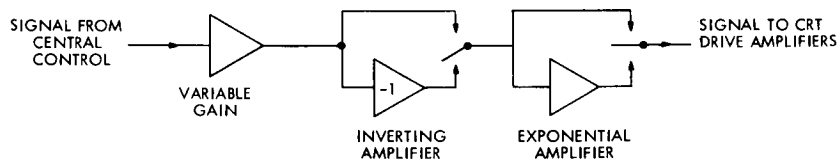
Fig. XIII-5. Scanner light feedback.

Fig. XIII-6. Display analog amplifiers.

It is important that the cathode-ray tube supply the same amount of light to each scanned point. The scheme for insuring this is shown in Fig. XIII-5. A pulse from the central control section initiates the intensification by putting a one-shot into a metastable state. While the tube is unblanked the current from the feedback PMT is integrated. The CRT is blanked when the integrated signal reaches a predetermined level or when the one-shot leaves the metastable state.

The schematic description of the amplifiers used to process the signal for the Display and for the Monitor is shown in Fig. XIII-6. Four transfer functions can be selected by interconnecting the various amplifier components. These are called: linear normal, linear inverted, exponential normal, and exponential inverted. The transfer functions of these modes are inverse to the corresponding transfer functions of the scanner.
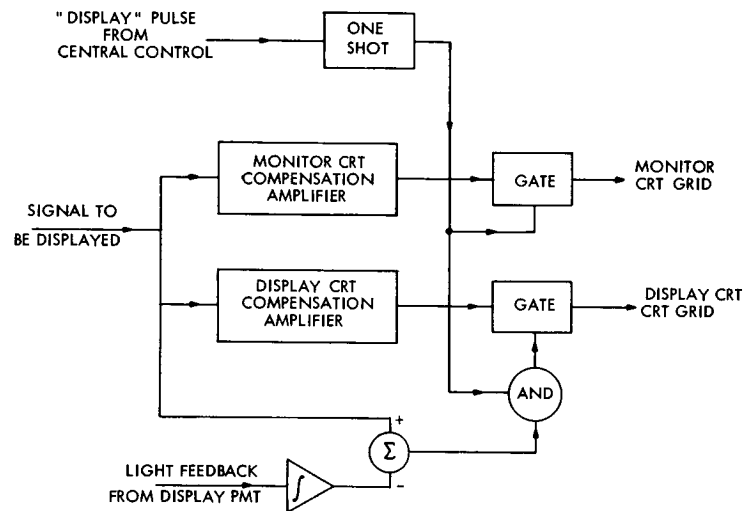

Fig. XIII-7. Display light feedback and monitor drive.

The circuits for driving the display and monitor cathode-ray tubes are given in Fig. XIII-7. The same signal is applied to the monitor and the display CRT. Since the transfer function of each tube is different, each tube has its own compensating amplifier. Both tubes are blanked most of the time. They are unblanked whenever this section receives a "DISPLAY" pulse from central control. This pulse triggers a one-shot. The monitor is unblanked as long as the one-shot is in the metastable state; the unblanking period for the display is controlled in the same way as in the scanner CRT. Light from the display unit is sampled by a PMT, the current from this tube is integrated, and this integral is compared with the signal to be presented on the display. The display tube is blanked when the one-shot leaves the metastable state or when the light feedback signal exceeds the input signal, whichever event occurs sooner. The controlling scheme is

intended to produce a linear relation between the exposing light from the face of the dis-
play tube and the signal at the input to Fig. XIII-7.

e. Deflection Section

The Deflection section contains flip-flop storage for the display coordinates,
D/A converters for changing the digital coordinates into analog voltages, and power
amplifiers for generating deflection
currents for the CRT. Control
amplifiers between the D/A con-
verters and the power amplifiers
enable one to change the corre-
spondence between the digital num-
ber and the deflection coordinates.
Gain and bias controls for these
amplifiers are located on the main
panel. The general arrangement
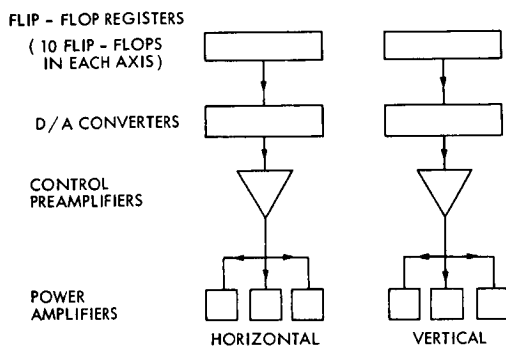of the Deflection system is shown
in Fig. XIII-8.

Fig. XIII-8. Deflection system symbolic diagram.

The flip-flop registers that hold
the display coordinates perform several functions. When the equipment is operated by
the computer, the registers can receive data directly from the computer. Under tape
operation, the registers operate as counters. A schematic diagram of these counters
is given in Fig. XIII-9. When SCAD is operated under local control, the scanning raster
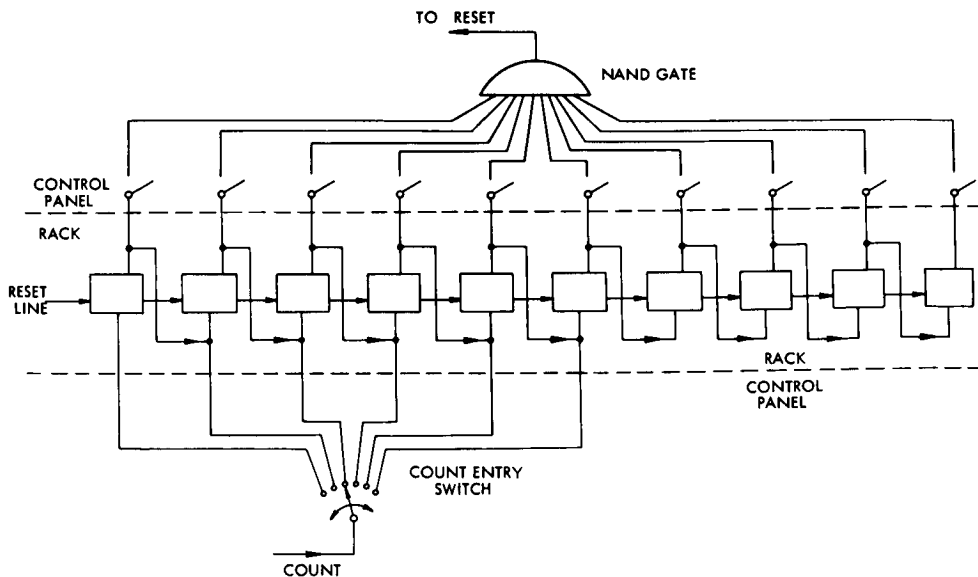

Fig. XIII-9. Counter interconnection for conventional noncomputer operation.

is determined by the counter settings. The two banks of counters are connected in tandem: the "clock pulses" are fed into the low-order end of the horizontal counter, and the "overflow" from the horizontal counter goes into the vertical counter. Both counters count up, and they start from zero. The raster is controlled by setting a number at which the counter is reset to zero, and by selecting the flip-flop which receives the "clock pulses." The last control, called "count entry," determines the fineness of the raster. Were the register allowed to go to a full count, one could change the raster resolution by a power of two with this control. Finer variations of the raster size are obtained with the above-mentioned reset control. The number at which the counter is reset is set up with a bank of toggle switches on the control panel.

The circuitry of the counter is such that the counter can be made to count either up or down. The direction of counter indexing is varied in the computer mode; under usual conditions of operation under local control, the direction of incrementing is fixed in both counters.

f. Central Control

The Central Control section performs a variety of functions. The operation of these is so interrelated that it is not convenient to represent this in block diagram form. We shall now describe some of the components contained in the Central Control section and some of the signals that are generated by it.

The Central Control section sequenced the operation of the other units that have been described. It generates command pulses to the Scanner, Display, etc. in response to signals from these units. For example, when the equipment is being used to play back tapes, the tape unit emits a pulse whenever a word of data is available. Central Control responds to this pulse by copying the data to the signal D/A converter and by displaying the data. After the Display indicates that the data have been displayed, Central Control issues a pulse that causes the counters to advance.

g. Test and Alignment Section

Several switches are connected to sundry test points in the SCAD. By operating these switches, one can select many of the important signals inside the SCAD and connect these to an oscilloscope. This facility is useful for aligning the signal from the scanner or to the display, and it is of some help in troubleshooting the equipment. The test section also contains some switches and circuitry that feed the deflection counter bits into the D/A converter, so that one can display a step hedge test signal.

h. Analog-to-Digital Converter

The analog signal produced by the scanner must be digitized before it can be sent to either the Tape Unit or the Computer. This is done with a bit-by-bit A-to-D converter.

The outputs are on parallel lines. The converter is preceded by a circuit that follows the input signal whenever the converter is quiescent, but holds the input voltage to the converter fixed when the unit is encoding this voltage.

i. Digital-to-Analog Converter

The digital signals coming from the Computer or from the Tape Unit are decoded into an analog voltage by the Signal Digital-to-Analog converter. This unit consists of a bank of 8 flip-flops which drives an 8-bit resistive ladder. The decoded voltage from the ladder is buffered with a unity-gain amplifier.

Each flip-flop in the register can be set from one of three sources. The data come either from the outputs of the A-to-D converter, from an external unit such as Tape or Computer or from the deflection counters.

Operating Modes

SCAD can be operated in one of several distinct modes. The mode is selected with the mode-control switch on the control panel. The internal programming for the mode controls is done through both the mode control switch and internal logic actuated by voltages that are generated with the switch.

Five principal modes are available. The modes fall into three types. The mode types and the actual modes are tabulated in Table XIII-1. The three types refer to different ways of obtaining data for the operation. In the two direct modes, the equipment takes the data obtained by digitizing the scanner output into the D/A converter and displays it. The two direct modes are called Direct Fast (DF) and Direct Slow (DS). The

Table XIII-1. SCAD operating modes.

Direct Modes:		
	1.	Direct Fast
	2.	Direct Slow
Tape Modes:		
	1.	Record
	2.	Playback
Computer Mode:		
	1.	Computer

difference between them will be explained below. The two different tape modes are called Record and Playback, and are used to display pictures recorded on tape and to record data from the scanner onto tape. In the Computer Mode, the equipment is operated by the computer.

We shall now discuss the modes in terms of the control signals that the Central Control section generates to operate the equipment in each mode. The sequencing section generates these commands in response to information that is received from subunits, such as the scanner and the converters, or from external devices such as the tape unit.

The command signals and "busy levels" used in SCAD are listed in Table XIII-2. We differentiate between a command, which is usually a pulse that initiates some sort of operation, and status levels, which indicate that a certain unit is performing its operation.

Table XIII-2. Command signals and "busy levels" used in SCAD.

Subunit	Command	Busy level	Remarks
A/D Converter	Convert	Converting	Convert Always follows end of Intensifying
D/A Converter	A/D to D/A	Copying	
	EXT to D/A	Copying	
	TEST to D/A	Copying	
	Reset D/A	(None)	Always occurs at end of displaying
Deflection Counters	Advance Hor	Deflecting	
	Advance Ver	Deflecting	
	Reset Hor	Deflecting	
	Reset Ver	Deflecting	
Display Unit	Display	Displaying	
Scanner	Intensify	Intensifying	
Tape Unit	Data Read		Signifies that the tape unit has some data waiting for SCAD
	Data Request		Tape unit calls for data from the scanner
	Data Ready		SCAD has data available for the tape unit
	Write EOR		SCAD requests the tape unit to write an EOR

An example of a command is DISPLAY: this is a pulse that causes the Display unit to display the data currently being put out by the D/A converter at the coordinates currently in the deflection registers. The "busy level" associated with this operation is called DISPLAYING. In the sequel we shall describe the chain of events that occurs in each mode, and comment on some of the items worthy of special note.

a. Matters Common to All Noncomputer Modes

In the noncomputer modes, the data format (which means the number of scanning lines per picture and the number of points per line) is determined by the counters. The number of points through which the horizontal or the vertical counter goes can be selected with the switches on the control panel. Under usual conditions, the count pulses enter the horizontal counter, and the overflow from the horizontal counter advances the vertical counter.

Whether or not the system is running is determined by the RUN flip-flop. This flip-flop is set by the start pushbutton. It can be reset by the stop pushbutton or by the overflow pulse from the vertical counter which will reset RUN if the single-multiple switch is in the single position; otherwise, the equipment will keep running.
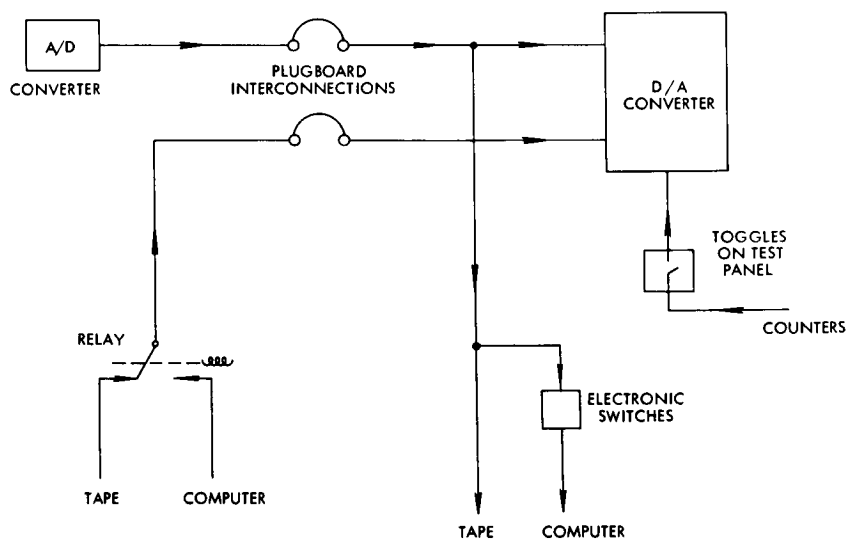
Fig. XIII-10. Digital signal-flow symbolic diagram.

The paths arranged for the flow of the digital data in the SCAD are shown in Fig. XIII-10. The noteworthy detail of this diagram is that all of the digital data go through a plugboard. This enables one to make some limited changes in the format of the data. For example, one may display, in the direct modes, pictures quantized to less than the 8 bits available from the A/D converter or

display the individual bits of the digitized picture.

b. Direct Fast Mode

The Direct Fast Mode is set up primarily for aligning and adjusting the Scanner and the Display. In this mode, the data from the scanner is transmitted directly to the display unit. The sequence of events in this mode is shown in Fig. XIII-11. In this figure, the name of a busy level followed by a bracketed E is used to represent the end of the activity associated with this level. For example, Displaying (E) represents the end of the displaying interval.
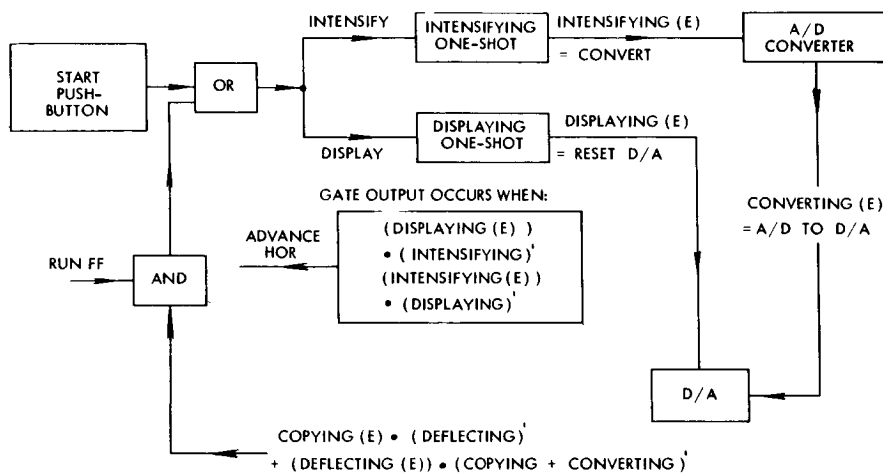
Fig. XIII-11. Direct fast-mode symbolic timing chart.

Please note that since the deflection coordinates of the scanner and the display unit are identical, the displayed picture in the Direct Fast mode does not appear at the same location as the original picture, but is "delayed" one dot. With the usual line-by-line scanning pattern, this misregistration is hardly noticeable.

c. Direct Slow Mode

The Direct Slow Mode can be used for the same purposes as the Direct Fast Mode. In the DS mode the data from the scanner is displayed at the coordinates from which it was obtained. The sequence of events for this mode is given in Fig. XIII-12.

d. Record Mode

In the Record Mode (refer to Fig. XII-13), the encoded data from the scanner are sent to the tape unit, which proceeds to record them. The timing in this mode is under control of the tape recorder: the tape recorder sends "Data Request" pulses to SCAD.
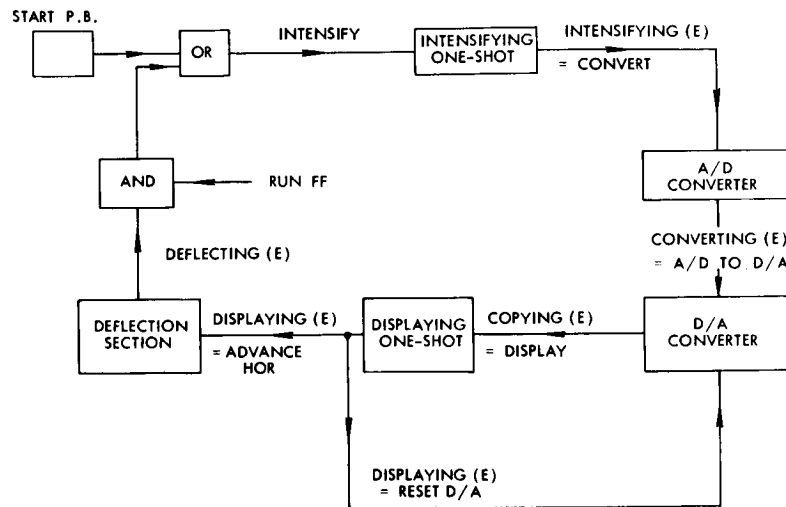
Fig. XIII-12. Direct slow-mode symbolic timing chart.

1. THE TAPE MOTION IS MAINTAINED
   BY THE RUN FF
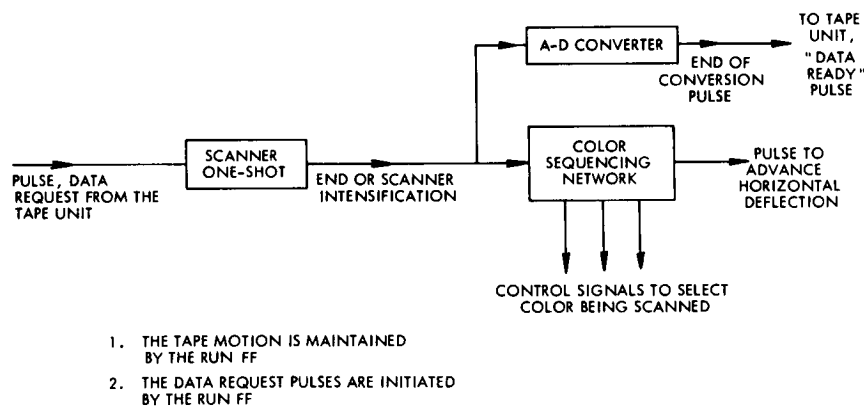2. THE DATA REQUEST PULSES ARE INITIATED
   BY THE RUN FF

Fig. XIII-13. Record-mode symbolic timing chart.

After a point has been interrogated and the brightness digitized, the SCAD replies with a "data ready" pulse. This pulse has to be within a certain time range after the Data Request.

The data on tape are blocked into records. The locations of the end of record gaps are determined by the SCAD: the tape recorder will insert a record gap whenever the SCAD sends it a Write EOR pulse. These pulses may be sent after every four, two or one scanning line. The number of lines per record is selected by a switch on the control panel.

The tape unit plays back the recorded data while it is recording, and these data are displayed. There is some delay between the recorded data and the time they are played back, so that the picture one sees on the monitor while the equipment is recording is

not a faithful replica of the signal actually on tape, but is useful primarily for monitoring the data.

e. Playback Mode

In the Playback Mode, the digital data from the tape unit are accepted by SCAD, converted to analog, and exhibited on both the Display and the Monitor. The timing chart for this operation is given in Fig. XIII-14 (the vertical advance and horizontal reset in the playback mode through the counters).
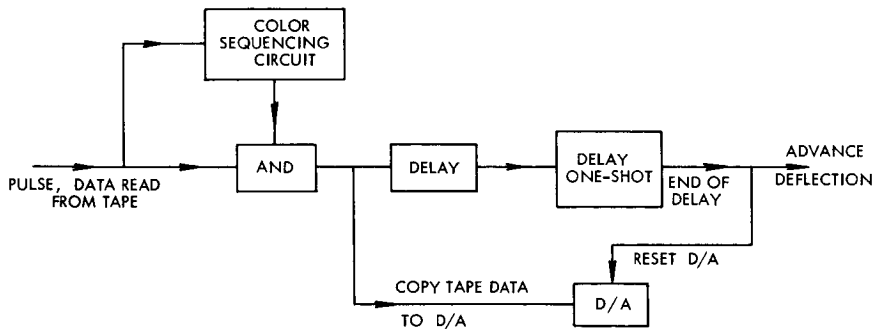
NOTES : 1   THE TAPE IS KEPT IN MOTION BY THE RUN
        2   THIS CHAIN OF EVENTS EXCEPT FOR DEFLECTION ADVANCEMENT ,
            IS ENABLED IN BOTH THE RECORD AND PLAYBACK MODES
        3   VERTICAL ADVANCE AND HORIZONTAL RESET ARE CAUSED BY THE END
            OF RECORD PULSE FROM THE TAPE UNIT WHEN THE EOR SWITCH IS UP

Fig. XIII-14. Playback symbolic timing chart.

The tape playback equipment detects the end of a tape record, and sends SCAD a pulse at this time. This pulse may be used as an alternate vertical advance and horizontal reset. The gate that permits this pulse to enter the counters is controlled by the Record Reset toggle on the front panel.

f. Color Operation

SCAD is capable of scanning and producing color pictures. The color signals are derived by splitting the light that passed through the transparency into three spectral bands, and measuring the amount of light in each band with a different PMT. Color pictures are obtained from the Display by photographing three separation pictures through three different color filters.

In the DF and DS modes, the color PMT that the equipment uses to obtain video is determined by switch setting on the control panel. In the record mode, the front-panel controls enable one to record any combination of the three color signals available; that is, one can record a monochrome recoding from each of the three PMT, or one of the

three possible two-color pictures, or a three-color picture. When a multicolor picture is recorded on tape, the numbers for the three chrominance components of each scanning point follow each other on the magnetic tape, i.e. , the three color signals are interleaved.

When this sort of polychrome tape is played back, the color signals alternate. To obtain a photograph from this, the equipment selects and displays the signal corresponding to only one color component. The controls for selecting the required signal are on the control panel.

g. Computer Operation

When the SCAD is operated in the Computer mode, it can be regarded as an input-output device for the computer. SCAD accepts three commands: DEFLECT, INTENSIFY, DISPLAY.

In the DEFLECT command, the SCAD accepts 18 bits from the computer, copies the more significant 9 into the vertical deflection register, and the less significant bits into the horizontal register. The INTENSIFY command is used to interrogate the scanner, and the DISPLAY command is used to present data on the Display unit and the Monitor. When an INTENSIFY command is given, SCAD intensifies the scanner at the current location, converts the brightness to digital, and sends these digital data to the computer. It next takes 4 bits from the computer, sets these into the deflection direction control register, and increments both the horizontal and vertical counters. Appropriate bits in this register enable the computer to step the deflection to any of the 8 adjacent points. The logic on the deflection direction control is interlocked so that the deflection cannot be made to stand in one point. The size of the increment made by the deflection is controlled by the count entry switches. The color that is sent to the computer is controlled by two bits that the SCAD obtains from the computer.

When SCAD receives a DISPLAY command, it takes eight bits from the computer, converts these to analog, and displays a spot of this brightness at the current coordinates. The deflection is stepped after the display in the same way as in the INTENSIFY command.

O. J. Tretiak

2. PICTURE TRANSMISSION BY PCM OVER A NOISY CHANNEL

In a previous report[1] a graphical method for minimizing the mean-square error when PCM is transmitted over a noisy channel was described, and the effect of using weighted PCM for transmitting pictures was shown. The improvement in subjective picture quality when weighted PCM is used is due to the fact that a large number of small errors in intensity is less disturbing to the viewer than a relatively small number of large errors.

Since that time, an investigation[2] was made of how minimizing another quantity than the mean-square error affects the quality of the received pictures, and simulations were carried out whereby the quantity $|\epsilon^w|$, with $w = 1.5$ and $w = 3$, was minimized. The results showed that the quality of the received pictures varies only slightly with the weighting power w, with an optimum value of $w = 2$ when the channel has initially an error probability of .001 and .005, and $w \approx 2.5$ if the initial error probability is .01.

A scheme, in which only the 5 most significant bits of the 6-bit pulse group were transmitted and the receiver made a guess on the least significant bit, was also tried. The quality of the pictures after transmission over the noisy channels was comparable to that of 5-bit originals. When only 4 bits were transmitted, however, and a guess was made on the two least significant bits, there were noticeable contours in the received pictures.

Some experiments in which unequal steps were used in quantizing the intensity of the picture points were also performed. If the quantization is done well, the same picture quality can be achieved by using only 5 bits instead of 6. The intervals were assigned numbers 0 to 31; these were transmitted over the channel, and the receiver determined the corresponding intensity from the received number. Because of the nonlinear mapping of the interval numbers into the intensity levels, the minimizing procedure used with the equally quantized pictures could not be applied to the intensity directly, but rather was applied to the numbers assigned to the intervals. Simulations were carried out by using weighting powers $w = 2$ and $w = 3$, and the picture quality was again increased considerably when weighted PCM was used.

Very noisy PCM channels, as assumed in this work, do not occur frequently in practice. An example is shown of how such a channel could arise when the rate is increased while transmitting over a bandlimited channel, thereby introducing errors because of intersymbol interference. Another simulation shows how weighted PCM improves the quality of the pictures when transmitted over such a channel.

This report is a summary of a Master's thesis carried out under the supervision of Professor T. S. Huang; the author wishes to acknowledge his guidance and many helpful suggestions throughout the work.

<div align="right">H. P. Hartmann</div>

References

1. H. P. Hartmann, "Reduction of the Output Noise Power in a Very Noisy PCM Channel," Quarterly Progress Report No. 81, Research Laboratory of Electronics, M.I.T., April 15, 1966, pp. 194-199.

2. H. P. Hartmann, "Picture Transmission by PCM over a Noisy Channel," S.M. Thesis Department of Electrical Engineering, M.I.T., August 1966.

3. DIFFERENTIAL PCM

Differential PCM is a digital communication system that transmits a quantized representation of the difference between a signal sample and its predicted value.[1-4] This system is shown in Fig. XIII-15. The input, s(t), is a discrete-time signal formed from samples of a bandlimited information signal (t may be regarded as an integer index). The quantizer is viewed as a device that adds noise n(t) to the quantizer input, d(t). The predicted value of s(t), r(t), is formed by adding x(t-1) to the sum of all previous x. It should be
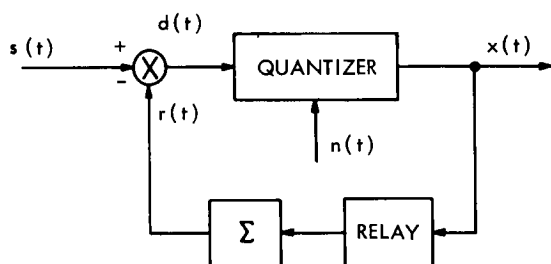
Fig. XIII-15. Differential PCM system.

noted that r(t) is also the system output and is formed at the receiving end in an identical manner. The following relations may be formed[5,6]:

$$r(t) = s(t-1) + n(t-1) \tag{1}$$

$$d(t) = \delta s(t) - n(t-1) \tag{2}$$

$$x(t) = \delta s(t) + \delta n(t), \tag{3}$$

where

$$\delta s(t) = s(t) - s(t-1)$$

$$\delta n(t) = n(t) - n(t-1).$$

Thus r(t) is the signal delayed by one unit plus the quantizing noise made the last time. The quantizer input, d(t), is the first difference of the information signal minus the delayed quantizing noise. The transmitted signal is the sum of the first differences of the information signal and the noise signal.

The noise power in a differential PCM system may be approximated by

$$\overline{n^2(t)} = \sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} [y_k - z]^2 P_\delta(z) \, dz, \tag{4}$$

where x_k and y_k are the kth decision and representation levels of the quantizer, $P_\delta(z)$ is the amplitude probability density of the first difference of the information signal, and N is the number of quantizing levels. Minimizing the noise power with respect to x_i and y_i leads to the following conditions on the differential PCM quantizer[7,8]:

$$x_i = \frac{y_i + y_{i-1}}{2} \tag{5}$$

$$\int_{x_i}^{x_{i+1}} [y_i - z] \, P_\delta(z) \, dz = 0. \tag{6}$$

The noise power in a standard PCM system of N quantizing levels is given by

$$\overline{n_p^2(t)} = \sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} [y_k - z]^2 \, P_\gamma(z) \, dz, \tag{7}$$

where $P_\gamma(z)$ is the amplitude probability density function of the information signal. It can be seen, by comparing of Eqs. 4 and 7, that if $P_\delta(z)$ is more peaked than $P_\gamma(z)$, the quantizer for a differential PCM system may be designed to yield a lower noise power than that for a standard PCM system. $P_\delta(z)$, in general, will be more peaked than $P_\gamma(z)$ for signals that are changing slowly with respect to the sampling frequency (i. e. , highly correlated signals with downward sloping spectra).

The improvement ratio of a differential PCM system over a standard PCM system is given by $\overline{n_p^2(t)}\big/\overline{n^2(t)}$. This ratio can be evaluated on a digital computer by using Eqs. 4 and 7. These equations can be used for an arbitrary set of quantizers.

Another comparison of differential PCM and standard PCM is provided[5,6] by

$$S/N = \frac{1}{2[1-R]} [K \, S/N_p - 1], \tag{8}$$

where

    S/N = signal-to-noise ratio of a differential PCM system

    S/N_p = signal-to-noise ratio of a standard PCM system

    K = the ratio of the differential PCM quantizer signal-to-noise ratio $[\overline{d^2(t)}\big/\overline{n^2(t)}]$ over the standard PCM signal-to-noise ratio

    R = the first displacement of the signal autocorrelation function divided by the signal power $[\overline{s(t) \, s(t-1)}\big/\overline{s^2(t)}]$.

Equation 8 illustrates the fact that the improvement of differential PCM over standard PCM is dependent on the degree of correlation of the input signal.

The transmitted signal of a differential PCM system is given by Eq. 3. When the sampling frequency is considered to be 1 sample/second, the power density spectrum

of $x(t)$, $G_x(\omega)$, is given by

$$G_x(\omega) = [G_s(\omega) + G_n(\omega)]\, T(\omega),$$

where $G_s(\omega)$ is the power density spectrum of $s(t)$, and $G_n(\omega)$ is the power density spectrum of $n(t)$. (It is assumed that $s(t)$ and $n(t)$ are uncorrelated.) $T(\omega)$ is defined as

$$T(\omega) = 2[1 - \cos \omega].$$

$T(\omega)$ is a sinusoid with a maximum value at one-half the sampling frequency, and thus tends to weight high frequencies much greater than low ones. If $G_s(\omega)$ is downward-sloping and $G_n(\omega)$ is flat (a characteristic of white noise), then the possibility exists that the noise power in a differential PCM system may be reduced at the receiving end by passing $s(t)$ through a linear filter that attenuates the higher frequencies.

A simulation of both differential PCM and standard PCM systems of 8, 16, and 32 quantizing levels was made on the I. B. M. 7094 digital computer located at the Computation Center, M. I. T. The input consisted of 256 samples of one scanning line of a photographic transparency. Optimum quantizers were approximated graphically by using Eqs. 5 and 6. Because of the small number of sample points used, the input signal lacked smooth statistical properties; however, the signal did exhibit the downward-sloping spectrum desirable for differential PCM. The power of the signal was 1.65×10^{-2} volt,^{2} and the correlation ratio was 0.921.

Table XIII-3 provides a brief summary of the results of this simulation.

Table XIII-3.   Comparison of differential PCM and standard PCM.

Quantizer Levels	PCM	Differential PCM			IR
	S/N_p	S/N_1	S/N_2	S/N_3	
8	56. 2	80. 8	101. 0	87. 8	1. 44
16	244. 0	520.	497.	523.	2. 13
32	1553	2460	3560	2570	1. 58

S/N_p and S/N_1 are the experimentally determined signal-to-noise ratios of the standard PCM and differential PCM systems, respectively. S/N_2 is the signal-to-noise ratio of the differential PCM system calculated by using Eq. 4. S/N_3 is the differential PCM signal-to-noise ratio calculated from Eq. 8, with the constant K experimentally determined. IR indicates the improvement ratio of differential PCM over standard PCM based on experimentally determined noise powers.

It can be seen by comparison of S/N_p with S/N_1 that the differential PCM system

yielded significant improvement over standard PCM for all cases. The improvement ratio was maximum at 16 levels, and decreased considerably at both 8 and 32 levels. Since there is some question about whether the quantizers that were used were truly optimum, these results should not be used to form any conclusions concerning the relationship between the number of quantizing levels and the improvement shown by differential PCM.

Equation 4 proved to be an accurate means of estimating differential PCM signal-to-noise ratio, with one major exception occurring at 32 levels. This happened because representation levels of the quantizer were situated at small isolated areas of the probability density $P_\delta(z)$. When this happened, no error power was contributed by Eq. 4 within this interval. Error will always be present in such an interval, however, since Eq. 2 states that the input to the quantizer never exactly equals the first difference signal.

Except for 8-level standard PCM, the noise in all systems was characteristically white (impulselike autocorrelation function and flat power density spectrum). The spectrum of the 8-level standard PCM system noise exhibited a downward-sloping spectrum similar to that of the input signal. A result like this is to be expected for coarsely quantized signals.

Finally, the improvement of differential PCM over standard PCM for this simulation cannot be considered to constitute a bandwidth reduction of one bit. This comparison should not obscure, however, the possibility of much greater bandwidth reduction resulting from the inherent subjective superiority of differential PCM as demonstrated for video signals.[2]

J. A. Newell

References

1. J. R. O'Neal, Jr. , "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals," Bell System Tech. J. 45, 689-721 (June 1966).

2. R. E. Graham, "Communication Theory Applied to Television Coding," ACFA Electronica, Vol. 2, pp. 333-343, 1957-1958.

3. B. M. Oliver, "Efficient Coding," Bell System Tech J. 31, 724-750 (July 1952).

4. J. A. Newell, "An Objective Study of Differential PCM," S. M. Thesis, Department of Electrical Engineering, M. I. T., 1966.

5. K. Nitadori, "Statistical Analysis of DPCM," Electronics and Communications in Japan (English Translation of J. Inst. Elec. Commun. Eng.-Jap.), Vol. 48, No. 2 February 1965.

6. R. A. McDonald, Unpublished work, Bell Telephone Laboratories, Inc.

7. J. Max, "Quantizing for Minimum Distortion," IRE Trans., Vol. IT-6, pp. 7-12, March 1960.

8. J. D. Bruce, "Optimum Quantization for General Error Criterion," Quarterly Progress Report No. 69, Research Laboratory of Electronics, M. I. T., April 15, 1963, pp. 135-141.

4. VARIABLE-VELOCITY DELTA MODULATION

Delta modulation[1] in its simplest form is shown in Fig. XIII-16. Incoming data are compared with the value in an accumulator (integrator), and a fixed-quantity "delta" is
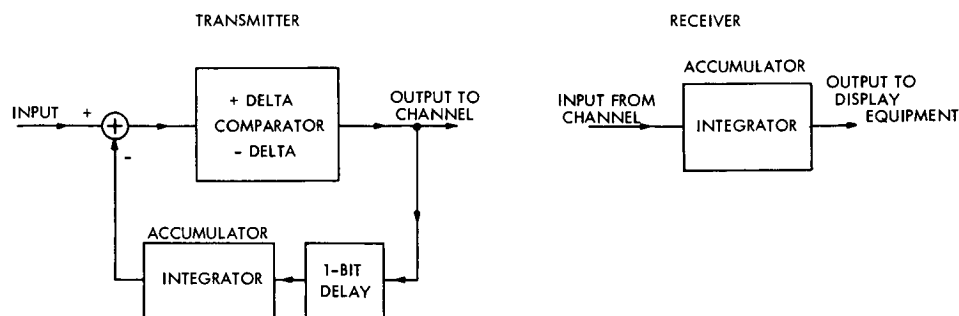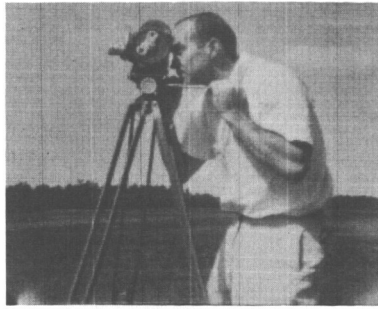

Fig. XIII-16. Simple delta modulation.

added to (or subtracted from) the accumulator so that it follows the input data. The output of the delta-mod transmitter is a string of bits that give the polarity information of each delta. At the receiver, the bits are decoded and added (or subtracted) to a similar accumulator whose value is the system output.

The cameraman picture original (Fig. XIII-17a) has been transmitted by a single-bit delta scheme, and the result is shown in Fig. XIII-17b. (All results shown were obtained by computer simulation.) Figure XIII-17b blurred because a too small delta was used. That is, the accumulator could not change to required brightness levels within a sufficiently few picture points. Figure XIII-17c shows the result of using a large delta. In this case the contours are sharp, but the range of different possible brightness is too restricted.

One popular solution to the problem of having too large or too small deltas is the use of two different sized deltas instead of only one.[2] Figure XIII-17d is an example of such a scheme, and the picture looks quite good; however, there has been a doubling of the transmitted information. In general, N bits/picture point will allow 2^{N-1} different sized deltas.

A new digital picture coding scheme has been invented which in some respects is similar to variable-velocity scanning[3] in analog T. V. systems. It has been named variable-velocity delta modulation[4] because the scan rate in the system is not constant. In all other delta-mod systems, the scan moves to the next picture point independently of whether or not the accumulator has changed enough to be a good approximation to the input data. This is the reason for the blurred contours of Fig. XIII-17b. The variable-velocity delta-mod transmitter keeps adding (or subtracting) deltas to the accumulator

(a)
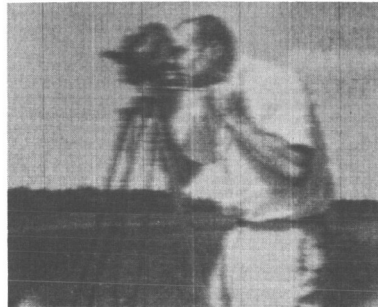
(b)
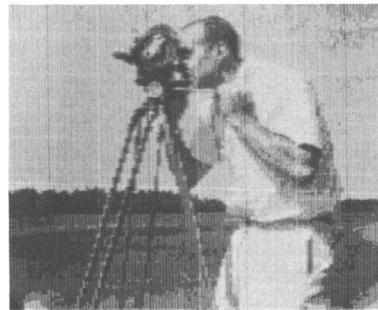
(c)

(d)

Fig. XIII-17.   Ordinary delta modulation.

until the input data value is equaled or the difference is within one delta, and only then does the scan move to the next picture point. A reversed polarity punctuation bit is added
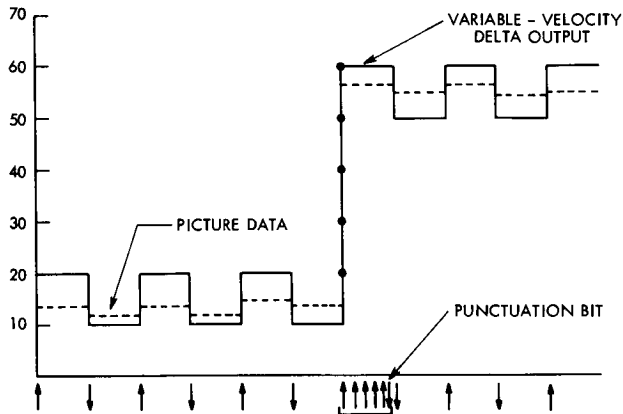

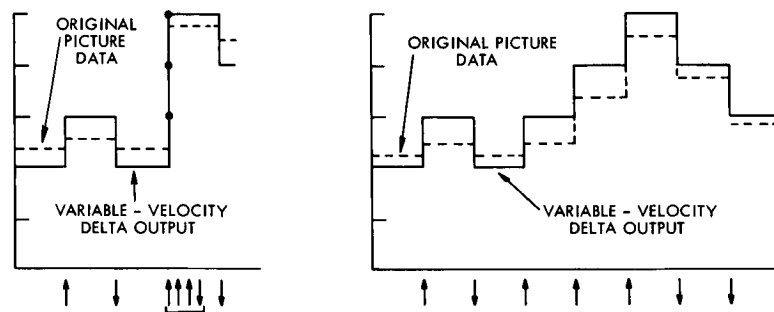Fig. XIII-18.   Variable-velocity delta modulation.

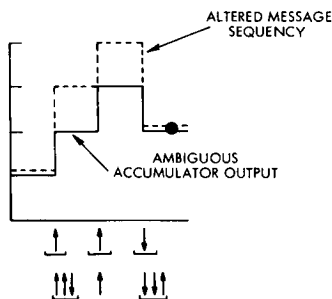Fig. XIII-19.   Example of ambiguous message sequence.

Fig. XIII-20.   Example of altering scheme.

to the end of strings of two or more identical polarity bits in order to help the receiver detect multi-delta picture points (see Fig. XIII-18). Note that when a picture has large regions of relatively constant brightness, adjacent picture-point differences will most

likely be within one delta, and therefore only one bit/picture point need be transmitted. The scheme is adaptive, since the number of bits/picture point is determined by the current needs, and therefore varies throughout the picture.

The receiver is designed to search for strings of identical polarity bits followed by a reversed polarity punctuation bit. For example, the five-in-a-row positive bits in Fig. XIII-18 are decoded as five deltas to be added to the receiver accumulator. It is possible for identical bit sequences to be transmitted from entirely different inputs (see Fig. XIII-19). To prevent such ambiguous sequences, it is necessary to test the bits to be transmitted for one of seven possible ambiguous cases. If one is found, then the message must be altered slightly so that the receiver will not make serious errors. The altering scheme introduces small errors, but they are hardly visible. One of the seven cases is shown in Fig. XIII-20. The top string of arrows represents the ambiguous message. Observe that the receiver would lump the first three bits together into the data

(a)

Fig. XIII-21. Variable-velocity delta
modulation.

(b)

for a single picture point.   The lower string of arrows, however, represents the bits that are actually transmitted, and the dotted line is what is seen at the receiver.   The remaining six ambiguous cases are given in the author's thesis.[4]

The cameraman, transmitted by variable-velocity delta-mod with a small delta used,

is shown in Fig. XIII-21a. The over-all average for the picture is 2.5 bits/picture point. When a larger delta is used, as shown in Fig. XIII-21b, the average drops to 1.75 bits/picture point. Note how much sharper the contours appear with variable-velocity delta-mod than with 2-bit delta-mod (i.e., 2 bits/picture point) as in Fig. XIII-17d.

S. D. Shoap

References

1. F. De Jager, "Deltamodulation, a Method of PCM Transmission Using the 1-Unit Code," Philips Research Report No. 7, December 1952, 442-466.

2. R. L. Remm, R. V. Cotten, and R. Strohmeyer, "Analysis and Implementation of a Delta Pictorial Encoding System," 1966 National Telemetering Conference Proceedings, pp. 27-34.

3. C. Cherry, M. H. Kubba, D. E. Pearson, and M. P. Barton, "An Experimental Study of the Possible Bandwidth Compression of Visual Image Signals," Proc. IEEE 51, 1507-1517 (1963).

4. S. D. Shoap, "Picture Encoding by Delta Modulation," S. M. Thesis, Department of Electrical Engineering, M. I. T., September 1966.

B. SENSORY AIDS

1. WORD-AT-A-TIME TACTILE DISPLAY

A device has been constructed, and partially evaluated, which permits the simultaneous presentation of a maximum of 8 Br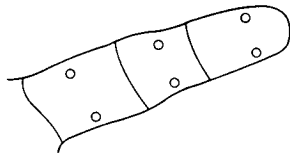aille-like cells to the fingers of both hands. An individual Braille cell is presented to each finger according to the stimulus locations shown in Fig. XIII-22. Here a character is very similar to the standard Braille character. Dots 1 and 4 stimulate the finger on the uppermost bone, dots 2 and 5 on the middle bone, and dots 3 and 6 on the lowermost bone next to the palm. With the hand placed palm down, dots 1, 2, and 3 stimulate the left side of the finger, while dots 4, 5, and 6 stimulate the right side.

Fig. XIII-22.   Stimulus locations.

As an individual Braille character requires only one finger, it is possible to present as many as 8 characters simultaneously (use of thumbs has been excluded for convenience). Therefore, words of 8 or less letters can be presented to the blind reader.

a. Description of the System

The system is capable of displaying blocks of information originating from punched paper tape, through tactile stimulation of the fingers, to a person at a rate that is variable from one block per 10 seconds to two blocks per second. Each block of information may consist of any amount of information from a single dot on a single finger to 6 dots on each of 8 fingers simultaneously.
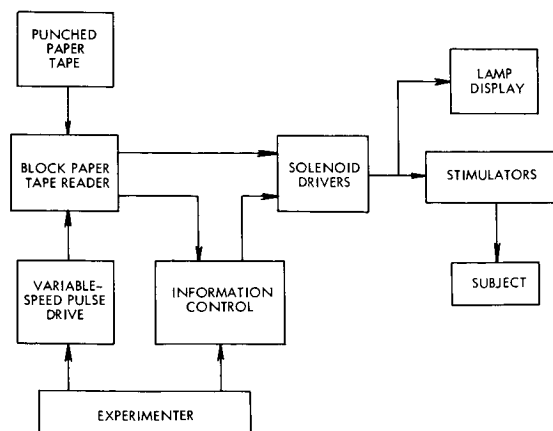
A block diagram of the system is shown in Fig. XIII-23. The block paper tape reader is capable of reading 12 rows of information from an eight-column tape, or 96 bits of information, on each cycle. The information is precoded onto punched paper tape by a PDP-1 computer program. The block paper tape reader is driven remotely by a variable-speed pulse drive which is controlled by the experimenter. The information control gates information to the solenoid drivers. Solenoids are used to provide the stimulus which is supplied to the fingers. The lamp display serves to

Fig. XIII-23.   The system.

moniter, for the experimenter, the drives applied to the solenoids in the stimulators.

b. Experiments and Conclusions

The objects of the experiments were, first, to determine whether or not Braille readers would be able to transfer their knowledge of Braille to the use of this system and hence be able to recognize individual stimulus patterns as letters and, second, to determine the reaction of the blind person to the system, by soliciting his opinions, criticisms, and suggestions for modification of the system.

Initially, each subject was familiarized with the system. The over-all operation was explained. The locations of the stimuli were indicated and the correspondence of the stimulus patterns on the finger to the Braille cell was emphasized.

Early experiments were designed to fix stimulus patterns with groups of letters in the subject's mind. The subject was presented multiple-letter patterns on his left hand, informed of their identity, and asked to verify them. Because of lack of time and slowness in progress, only the left hand was used.

In later experiments a pattern was presented and the subject asked to identify it, being corrected, if necessary, immediately afterward. This was done with single-letter patterns, two-letter patterns, and common three-letter words.

The test described here consisted of a random finger test, summarized in Figs. XIII-24 and XIII-25. The purpose of this test was to determine the relative ability
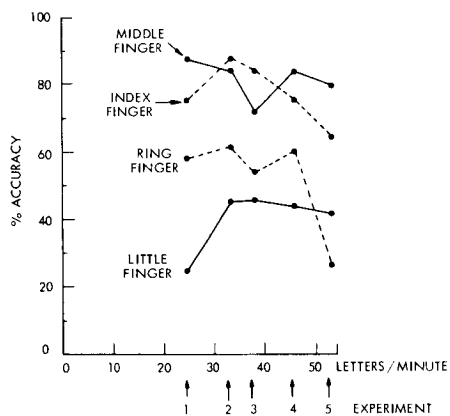

Fig. XIII-24.   Random finger test. Each point represents an average of 25 data. (Subject: S_1.)

of each of the fingers to recognize single characters. For each experiment in this test 100 letters were presented, one at a time, in random fashion, so that each finger had exactly 25 letters presented to it. At any given time during an experiment only one finger had a letter presented to it. These experiments were performed for each of the

letter rates given below. In each experiment the same letters were used but in different order.

1. 25 letters per minute
2. 33 letters per minute
3. 38 letters per minute
4. 46 letters per minute
5. 53 letters per minute.

(The fifth experiment was not performed on subject S_2.) The results of this test support the previous belief that the lit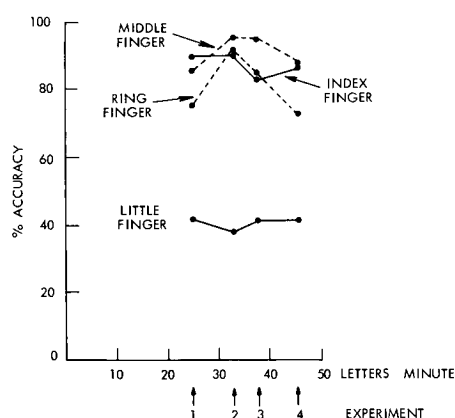tle finger is least able to perceive the stimulus. There are insufficient data here to determine which finger, if any, is best able to perceive stimulus. The weakness of the little finger gives a hint, however, that when a word is to be presented to both hands jointly, it may be advantageous to center the word about the two index fingers when it requires less than all 8 fingers.

A more complete description of the other tests, as well as the opinions and suggestions of the subjects, may be found in the author's thesis.[1]

In conclusion, it appears to be correct to say that there is considerable transfer of knowledge of the Braille code, as normally read, by the subjects to the recognition of individual patterns as displayed by this system. Certainly, more extensive tests are required before a complete evaluation of this system can be made.

J. A. Williams

Fig. XIII-25. Random finger test. Each point represents an average of 25 data. (Subject: S_2.)

References

1. J. A. Williams, "Word-at-a-Time Tactile Display," S. M. Thesis, Department of Electrical Engineering, M. I. T. , May 1966.

2. METHOD FOR STORAGE OF BRAILLE

Research is under way at the Massachusetts Institute of Technology, and elsewhere, on machines to identify (recognize) printed characters, independently of type size, style and font. Such machines, which are approaching practicability, are intended for use in reading machines for the blind. Most workers have attempted to provide an output to

the blind user in the form of spelled speech or phoneme strings,[1,2] in spite of the fact that with present technology such a machine would be prohibitive in terms of size and price to a typical, prospective user.

While it is now possible to build a machine that will read print and produce it as Braille on-line, it will be some time before such a machine can be made portable, and in fact, it may never be economically feasible to do so. A centrally located reading machine providing embossed Braille books would not be satisfactory; Braille books are heavy and bulky, and as reading machines become widespread, and the amount of reading material available to blind people approaches that available to sighted people, blind people with only moderately sized libraries will encounter severe storage problems.

As a solution to this problem, this report suggests that the output of the reading machine be in the form of magnetic tape, which, in turn, is converted to Braille. The approach is intended for use with a central facility for translating printed material to tape. The user would store the tape, convert it to Braille as needed on a portable unit, and destroy the Braille after use. A scheme has been developed with the following provisions:

1. compactness of storage on tape comparable to that of print, in terms of size and weight;

2. conversion from magnetic tape to Braille on inexpensive, portable equipment;

3. conversion from magnetic tape to Braille at continuously variable word rates, ecompassing all speeds at which Braille is read; and

4. an error check, and reasonably low error rates.

The magnetic tape record/playback unit was designed for use with braillers that require input information at the word speed at which the user chooses to read; accordingly, the magnetic tape unit must provide output at continuously variable word rates. Hence the tape must be either intermittently clutched or drawn steadily across the heads at continuously variable rates. It was considered impractical to build a portable unit that drives tape intermittently, and so the alternative, that of variable tape speeds, was selected.

Tape data systems almost always use saturation recording techniques and some form of frequency or phase modulation, because of amplitude variations in the so-called linear region of the tape.[3,4] The present requirement of continuously variable tape speeds rules out phase- and frequency-modulation systems because both would require critical settings of a playback oscillator or filter. Amplitude-modulation systems have the overwhelming advantage of being relatively insensitive to tape speed, since it is easy to compensate for the fact that over-all amplitudes vary linearly with tape speed. As for amplitude variations caused by tape imperfections, their effect can be reduced to a tolerable level by using simple redundancy. Since the final link in the system is to be a human possessed of considerable error-correcting ability, errors are not as undesirable

as they would be, for example, in a data-processing system. It was estimated that one error per 100 or 200 characters would not be objectionable,[5] and this order of error rate was achieved with the amplitude-modulation scheme described here.

The AM technique has two further advantages, in that the scheme lends itself to realization with a minimum of hardware in the playback unit, and hence facilitates portability. To the same end, it developed that any small tape recorder that could handle speed intelligibly would suffice for recording and playing back Braille, when appropriately modified. The resulting tape unit is potentially portable and inexpensive.

A ternary, or three-level, code is employed, and all signals are recorded at a single frequency. (Effective frequency on playback varies with playback speed.) The three possible signals on the tape are a "high" bit, consisting of 8 cycles at a relatively high amplitude, a "low" bit, consisting of 8 cycles at some lower amplitude, and a "null" bit, consisting of 8 periods of zero amplitude. Each Braille cell is encoded in 9 bits: the first bit is always high, the next 6 bits are high or low, in such a way as to represent dots and their absences, respectively, in the Braille cell, the eighth bit makes the over-all parity odd, and the ninth bit is a null bit, used for synchronization.
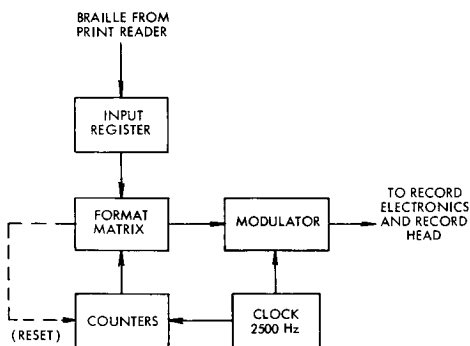
The record section is entirely straightforward, and can be seen in Fig. XIII-26. It serves to generate the format described above.

The playback section is somewhat more involved (Fig. XIII-27). Two separate counter chains are employed: one is reset by the null between characters and, by counting all cycles ("pulses"), keeps track of which bit in the Braille cell is under the playback head. The other counter registers only high-amplitude cycles, and so judges each bit to be high or low. If the majority of the 8 cycles in a bit is received as high, the bit
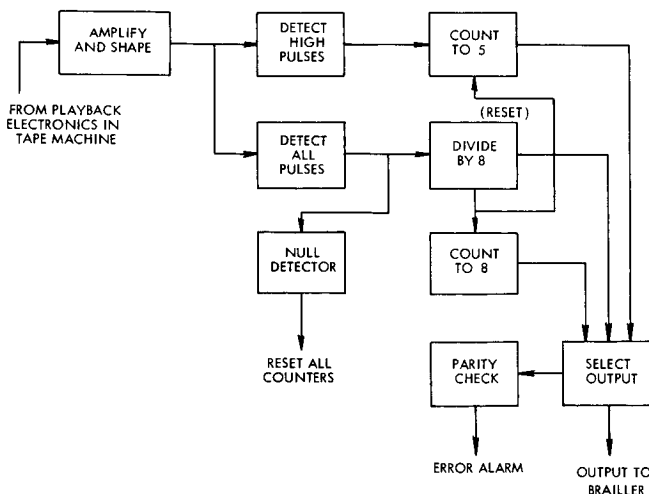

Fig. XIII-26. Record section.

Fig. XIII-27. Playback section.

is counted as high, otherwise, as low. This information is passed on to the brailler, together with an audible alarm that the playback parity is not odd.

It is clear that every cycle need not register perfectly in order that a character be correctly interpreted. For example, the first counter can gain or lose up to 4 cycles per character, and still cause perfect output. The second counter can wrongly judge (as high or low) up to 4 cycles per bit, and still produce perfect output. Furthermore, the use of the null bit between characters prevents the carrying over of wrong counts from one character to the next. The use of 8 cycles to encode each bit offers considerable protection against noise, while maintaining a reasonable words-per-reel density.

A prototype was constructed to test the feasibility of these ideas. Sample material was recorded at 7 1/2 inches per second, with a "carrier" frequency of 2500 Hz. This corresponds to 416 words per minute, which is faster than Braille reading rates. The system was intended to be continuously variable down to 0.18 inch per second, which gives a playback carrier frequency of 60 Hz and a bottom reading rate of 10 words per minute. With the encoding scheme described here, and 8 tracks consecutively of standard, 1-mil, one-quarter inch audio recording tape, one seven-inch reel stores 140,000 English words. This words-per-volume ratio is superior to that of hard-cover books, and is comparable to that of thin-paper paperback books. Experimental error rates were approximately one character in 200, better than current print readers by an approximate factor of twenty.

M. B. Lazarus

References

1. H. Freiberger and E. F. Murphy, "Reading Machines for the Blind," IRE Trans. Vol. HFE-2, No. 1, pp. 8-19, March 1961.

2. P. W. Nye, "Reading Aids for Blind People," Med. Electronics Biol. Eng., Vol. 2, No. 3, July 1964.

3. G. L. Davies, Magnetic Tape Instrumentation (McGraw Hill Book Company, New York, 1961).

4. H. G. M. Spratt, Magnetic Tape Recording (Heywood and Company, London, 1958).

5. G. A. Miller and E. A. Friedman, "Reconstruction of Multilated Texts," Inform. Contr. 1, pp. 38-55 (September 1957).

XIV. COMMUNICATIONS BIOPHYSICS[*]

Academic and Research Staff

Prof. D. B. Geselowitz[†]　　Prof. T. F. Weiss[‡]　　　　Dr. M. Nomoto[***]
Prof. P. R. Gray　　　　　　Dr. J. S. Barlow[††]　　　　　Dr. R. R. Pfeiffer[‡]
Prof. H. B. Lee　　　　　　　Dr. A. W. B. Cunningham　　Dr. R. Rojas-Corona
Prof. W. T. Peake[‡]　　　　　N. I. Durlach　　　　　　　　Dr. G. F. Songster
Prof. W. A. Rosenblith　　　Dr. H. Fischler[‡‡]　　　　　　R. M. Brown[‡]
Prof. W. M. Siebert　　　　　Dr. R. D. Hall　　　　　　　　A. H. Crist[‡]
Prof. R. Suzuki[**]　　　　　　Dr. N. Y. S. Kiang[‡]　　　　D. P. Langbein[‡]

Graduate Students

J. A. Anderson　　　R. E. Greenwood　　E. C. Moxon
G. von Bismarck　　J. J. Guinan, Jr.　　M. Nahvi
L. D. Braida　　　　D. K. Hartline　　　P. H. O'Lague
S. K. Burns　　　　L. K. Krakauer　　　D. J-M. Poussart
H. S. Colburn　　　R. G. Mark　　　　　M. B. Sachs
J. E. Evans　　　　E. G. Merrill　　　　I. H. Thomae
J. A. Freeman　　　P. J. Metz III　　　M. L. Wiederhold
　　　　　　　　　　D. C. Milne

A. AUDITORY NERVE FIBER RESPONSES TO TWO-TONE STIMULI

[This report gives a brief summary of a Ph.D. thesis of the same title, submitted to the Department of Electrical Engineering, M.I.T., August 1966. Further details and results may be found in the thesis itself.]

Electrophysiological investigations of auditory nerve fiber responses thus far have been limited largely to what have been called "simple" acoustic stimuli: clicks, tones, and tone bursts. Some recent findings[1,2] have revealed rather complicated interactions in the responses to more complex stimuli and have indicated that extension of results from simple to complex stimuli may not be straightforward. There have been few

quantitative data relating to these interactions, however. In this study we have attempted
to describe quantitatively the responses of auditory nerve fibers to somewhat more com-
plex stimuli, sums of two tones. We have chosen to describe the fiber responses by a
single variable, rate of spike discharge.

We have used a sweep-frequency technique[3] to construct iso-rate contours for
responses to one- and two-tone stimuli. An iso-rate contour is the locus of points in
the frequency-level plane corresponding to stimuli that evoke a specified rate of response
from a fiber. Such contours provide a convenient means for illustrating general charac-
teristics of the dependence of rate on stimulus frequency and level. When measuring
rate as a function of the frequency of one of two tones, we presented that tone as a sweep-
frequency signal. If a second tone was present, it was presented as a continuous tone at
the fiber's characteristic frequency (CTCF). Figure XIV-1 shows iso-rate contours for
a typical fiber for both one- and two-tone stimuli. In this figure we have specified stim-
ulus levels in terms of peak-to-peak stapes displacement. In doing so, we hope to iso-
late effects produced by the cochlea from effects produced by the frequency dependence
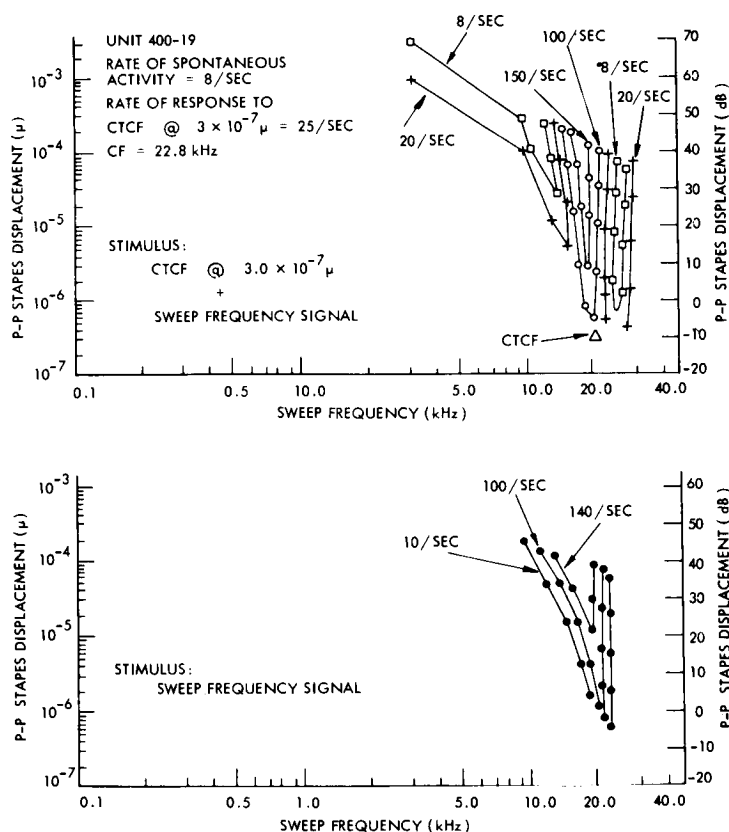of the middle-ear transmission.[4]

Fig. XIV-1. Iso-rate contours.

For one-tone stimuli (lower portion of Fig. XIV-1) we find a limited region in the stimulus plane within which a fiber's rate of discharge is increased above its spontaneous level. For each fiber the "excitatory response area" is defined to be that area in the stimulus plane bounded below by the iso-rate contour corresponding to a rate 20 per cent greater than the spontaneous rate. From Fig. XIV-1 it is evident that all stimuli lying within a fiber's excitatory response area evoke a rate of response at least 20 per cent above the spontaneous rate.

In the two-tone stimulus situation (CTCF plus sweep-frequency signal – upper section of Fig. XIV-1) the fiber exhibits a region in the sweep-frequency stimulus plane in which the effect of the sweep-frequency signal is to increase the rate of discharge above its value for the CTCF presented alone. To the right and left of this region are regions in which the effect of the sweep-frequency signal is to reduce the rate below its value for the CTCF alone. For any CTCF level, we define the "inhibitory response area" to be the region in the sweep-frequency stimulus plane bounded below by the iso-rate contours corresponding to a rate equal to 80 per cent of the rate to the CTCF alone. Any tone with stimulus parameters within the inhibitory response area causes a reduction in the rate of response to the CTCF of at least 20 per cent. We call the reduction in the
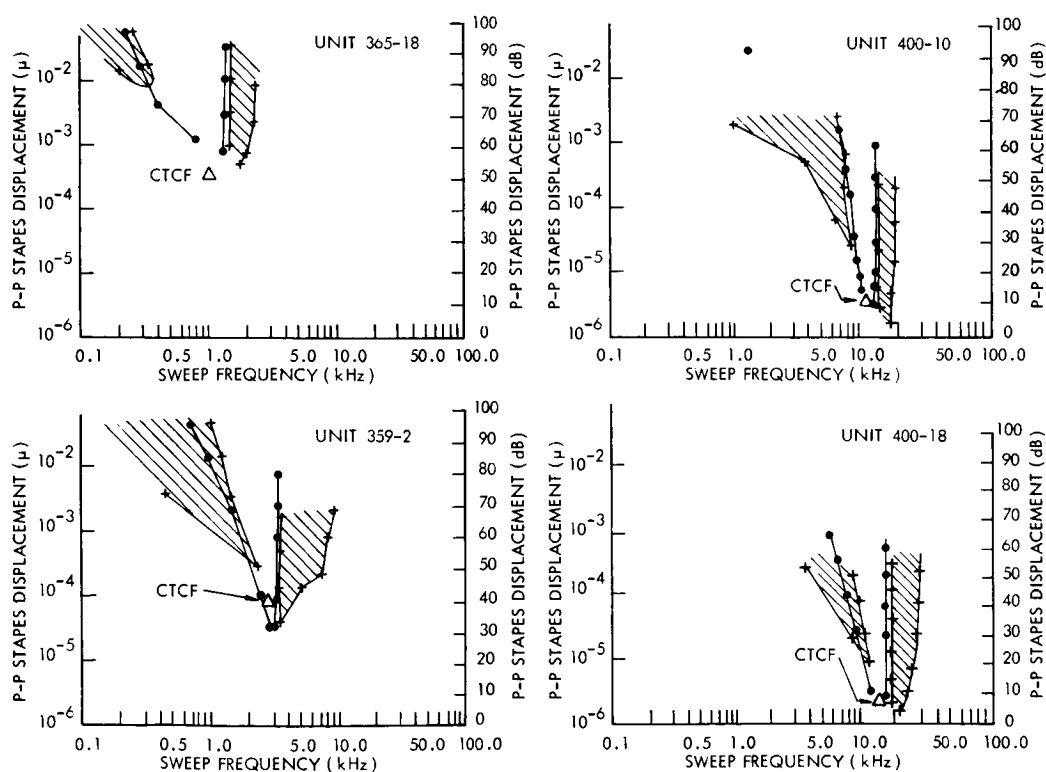

Fig. XIV-2. Excitatory and inhibitory response areas for four fibers.

fiber's rate of discharge to a CTCF by the presentation of a second tone "two-tone inhibition." All second tones with stimulus parameters within the inhibitory response area produce two-tone inhibition. We have observed two-tone inhibition for more than 300 fibers; we have never failed to find two-tone inhibition for any fiber when we were able to look for it systematically.

Figure XIV-2 shows excitatory and inhibitory response areas for four fibers covering a broad range of characteristic frequencies. These areas are typical of more than 200 fibers. The fibers all have inhibitory areas on both sides of their excitatory response areas. We have found that the extent of the inhibitory areas depends on the level of CTCF employed. An increase in CTCF level causes an upward shift in the boundaries of the inhibitory areas in the sweep-frequency stimulus plane. As is illustrated by Unit 359-2 (Fig. XIV-2), if the level of the CTCF is low enough, the inhibitory area above the CF extends down to levels very close to the tip of the excitatory response area. Higher levels are needed, however, to produce two-tone inhibition at frequencies below fiber CF. The minimum levels necessary to reduce rate by a given amount at frequencies below the CF are at least 20 dB greater than the minimum levels of stapes displacement needed at frequencies above the CF.
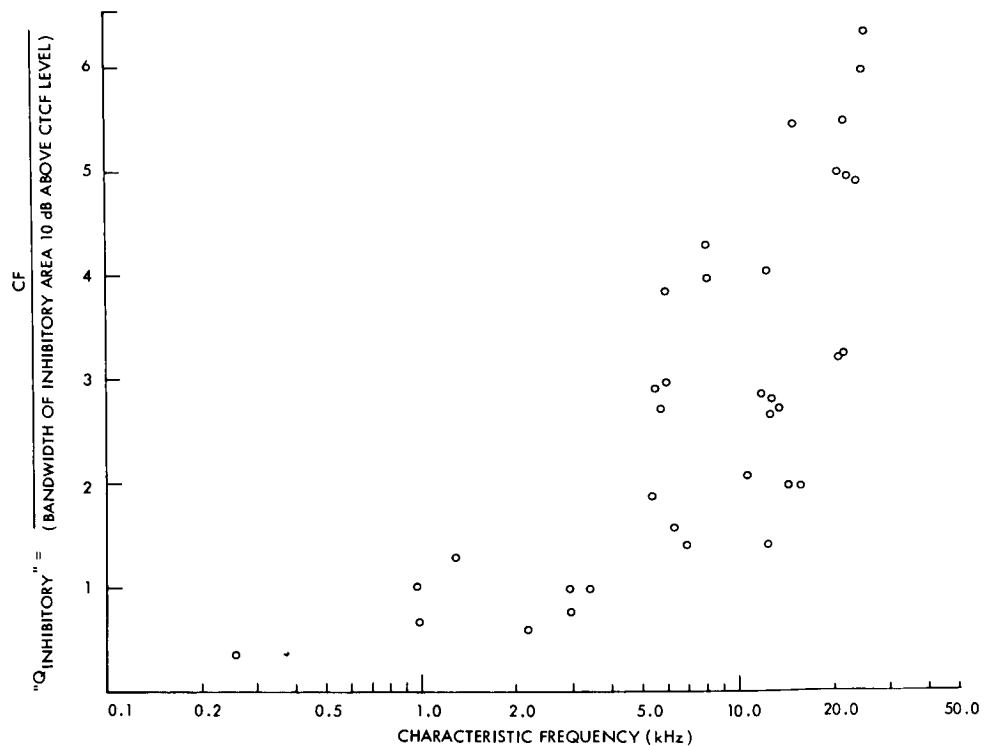

Fig. XIV-3. "Q_{INH}" as a function of CF.

An appropriately defined "Q" can be taken as a measure of the extent of the inhibitory areas in the sweep-frequency stimulus frequency dimension. We have adopted the following definition of the "Q" of an inhibitory area: Q_{INH} = CF/bandwidth of the inhibitory area 10 dB above the CTCF level. Figure XIV-3 shows a scatter plot of Q_{INH} versus fiber CF for 35 fibers. The dependence of these inhibitory Q's on CF is similar to the dependence of the tuning curve Q's given by Kiang,[1] in that both the tuning curve Q's and inhibitory Q's are roughly constant for CF below 2 kHz, and for higher frequencies both Q's increase with increasing CF.

In order to collect data with which we could test mathematical descriptions of the two-responses, we used tone-burst stimuli and measured the discharge rate after the transient at the tone-burst onset. Using this method, we developed the following expression to describe the relationship between discharge rate and the stimulus parameters of the two-tones. (Our data are restricted to the case in which one tone is at fiber CF.)

$$r(P_{CF}, f_C, P_2, f_2) = R_{SP} + R(P_{CF}, f_C)\, g\left(\frac{P_2}{P_{CF}}, f_C, f_2\right) + R(P_2, f_2)\, g\left(\frac{P_{CF}}{P_2}, f_2, f_C\right),$$

where

$r(P_{CF}, f_C, P_2, f_2)$ = spike discharge rate as a function of the two-tone stimulus parameters

P_{CF} = sound-pressure level of the tone at the characteristic frequency

f_C = characteristic frequency

P_2 = sound-pressure level of the second tone

f_2 = frequency of the second tone

$R_{SP} = r(0, f_C, 0, f_2)$ = spontaneous discharge rate

$R(P_{CF}, f_C) = r(P_{CF}, f_C, 0, f_2) - R_{SP}$ = "driven part" of the rate to one tone alone

$g\left(\dfrac{P_2}{P_{CF}}, f_C, f_2\right)$ = "inhibitory multiplier."

The function $R(P, f)$ is a monotone nondecreasing function of P. The inhibitory multiplier $g\left(\dfrac{P_2}{P_{CF}}, f_C, f_2\right)$ is 1 for a small value of P_2/P_{CF} and is a monotone decreasing function of P_2/P_{CF}. This mathematical form has been adequate to "fit" all data that we have obtained, although we lack sufficient data for certain conditions. We have suggested how this result might be extended to more general two-tone stimuli (neither tone at the CF) and to the sum of an arbitrary number of tones.

M. B. Sachs

References

1. N. Y. S. Kiang, Discharge Patterns of Single Fibers in the Cat's Auditory Nerve, M.I.T. Research Monograph No. 35 (the M.I.T. Press, Cambridge, Mass., 1965).

2. M. Nomoto, N. Suga, and Y. Katsuki, "Discharge Patterns and Inhibition of Primary Auditory Nerve Fibers in the Monkey," J. Neurophysiol. 27, 768-787 (1964).

3. R. R. Pfeiffer and A. H. Crist, "A Method for Studying Response and Inhibition Areas of Some Single Units in the Auditory System," Quarterly Progress Report No. 80, Research Laboratory of Electronics, M.I.T., January 15, 1966, pp. 232-238.

4. J. J. Guinan and W. T. Peake, "Motion of Middle-Ear Bones," Quarterly Progress Report No. 74, Research Laboratory of Electronics, M.I.T., July 15, 1964, pp. 219-221.

XV. NEUROPHYSIOLOGY[*]

Academic and Research Staff

A. STABILITY OF NETWORKS WITH LOOPS[†]

1. State Transition Matrix of a Neural Network

Let $\mathcal{N}$ represent a network of N formal neurons with interacting afferents,[1] i.e., neurons which are capable of computing any Boolian function of their inputs, that contain M external inputs, and are organized by means of internal loops. We denote by $x_1(t)$, $x_2(t)$ ... $x_M(t)$ the external inputs at time t, whereas $y_1(t)$, $y_2(t)$ ... $y_N(t)$ denotes the outputs at time t, which are also regarded as the state $\vec{S}(t)$ of the network at that time. We may therefore write,

$$\vec{S}(t) = (y_1(t), y_2(t), \dots y_N(t)).\tag{1}$$

The number of network states is 2^N, which we denote by $\vec{S}_i (i=1, 2, \dots 2^N)$. We denote by $\vec{X}(t)$ the input configuration $x_1(t)$, $x_2(t)$, ... $x_M(t)$ at time t, of which there exist 2^M different input configurations, $\vec{X}_m (m=1, 2, \dots 2^M)$.

We define the network $\mathcal{N}$ by a set of N Boolian functions of the form:

$$y_1(t) = f_1[x_1(t-1), \dots x_M(t-1); y_1(t-1), \dots y_N(t-1)]$$

$$y_2(t) = f_2[x_1(t-1), \dots x_M(t-1); y_1(t-1), \dots y_N(t-1)]$$

$$\vdots \tag{2}$$

$$y_N(t) = f_N[x_1(t-1), \dots x_M(t-1); y_1(t-1), \dots y_N(t-1)].$$

Equations 2 may be written as

$$y_1(t) = f_1[\vec{X}(t-1); \vec{S}(t-1)]$$

$$y_2(t) = f_2[\vec{X}(t-1); \vec{S}(t-1)]$$

$$\vdots$$

$$y_N(t) = f_N[\vec{X}(t-1); \vec{S}(t-1)].$$

(3)

For a particular configuration, $\vec{X}_m$, of the inputs, Eqs. 3 are Boolian functions of $y_1, y_2, \ldots y_N$, i.e.,

$$y_1(t) = f_1[\vec{X}_m; \vec{S}(t-1)]$$

$$y_2(t) = f_2[\vec{X}_m; \vec{S}(t-1)]$$

$$\vdots$$

$$y_N(t) = f_N[\vec{X}_m; \vec{S}(t-1)].$$

(4)

For each value of $\vec{S}(t-1) = \vec{S}_i$, we obtain a new state $(y_1(t), y_2(t), \ldots y_N(t)) = \vec{S}_j$. For each input $\vec{X} = \vec{X}_m$, we wish to consider the state transition matrix $\mathscr{M}(\vec{X}_m)$, i.e., the Boolian matrix of 2^N rows and columns in which the $\mathscr{M}(\vec{X}_m)_{ij}$ term is 1 if the network goes from the state $\vec{S}_i$ to the state $\vec{S}_j$ under the input $\vec{X}_m$, and 0 otherwise. Therefore, these matrices $\mathscr{M}(\vec{X}_m)$ have one and only one 1 in each row. If $(\alpha, \beta, \ldots \nu)$ are the components of $\vec{S}_j$, i.e., they constitute the string of zeros and ones that define $\vec{S}_j$, $\mathscr{M}(\vec{X})_{ij}$ may be written

$$\mathscr{M}(\vec{X})_{ij} = f_1^\alpha(\vec{X}; \vec{S}_i) \cdot f_2^\beta(\vec{X}; \vec{S}_i) \cdot \ldots \cdot f_N^\nu(\vec{X}; \vec{S}_i)$$

(5)

following the convention in which

$$f_n^O(\vec{X}; \vec{S}_i) = \overline{f_n(\vec{X}; \vec{S}_i)} \qquad \text{(negation)}$$

$$f_n^1(\vec{X}; \vec{S}_i) = f_n(\vec{X}; \vec{S}_i).$$

(6)

Example 1. Consider the network of Fig. XV-1, for which the functions f_1 and f_2 are given by

$$f_1 = x_1 \bar{y}_1 \bar{y}_2 + \bar{x}_1 y_1 \bar{y}_2 + \bar{x}_1 \bar{y}_1 y_2$$

$$f_2 = x_2 \bar{y}_1 + \bar{x}_2 y_1.$$
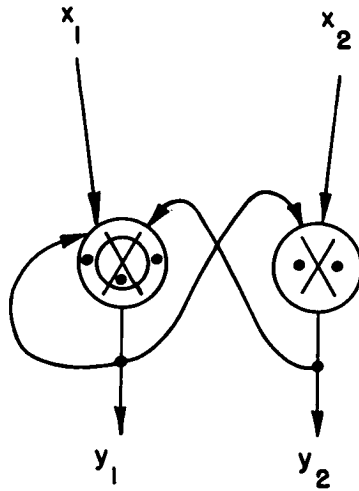
The state transition matrix, $\mathscr{M}(\vec{X})$, is

Fig. XV-1.

Fig. XV-2.
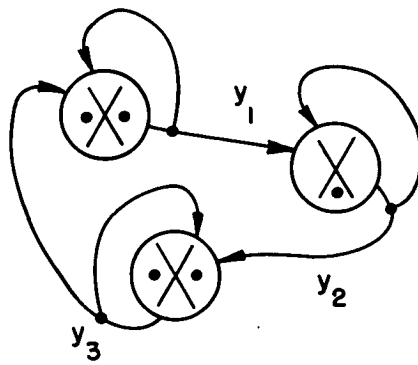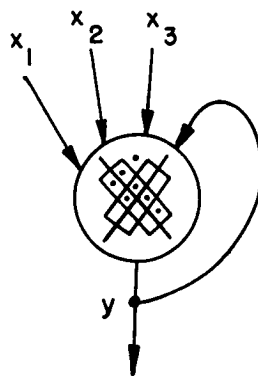

Fig. XV-3.

$$\vec{S}_j$$

	00	01	10	11
00	$\bar{x}_1 \bar{x}_2$	$\bar{x}_1 x_2$	$x_1 \bar{x}_2$	$x_1 x_2$
01	$x_1 \bar{x}_2$	$x_1 x_2$	$\bar{x}_1 \bar{x}_2$	$\bar{x}_1 x_2$
10	$x_1 x_2$	$x_1 \bar{x}_2$	$\bar{x}_1 x_2$	$\bar{x}_1 \bar{x}_2$
11	x_2	$\bar{x}_2$	0	0

$\vec{S}_i$ labels the rows.

For example, the input $\vec{X}_m = (0, 0)$, gives $\mathscr{M}(0, 0)$,

$$\mathscr{M}(0,0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

which means that, under the input $(0, 0)$, the transitions of states are

$(0, 0) \rightarrow (0, 0)$

$(0, 1) \rightarrow (1, 0)$

$(1, 0) \rightarrow (0, 1)$

$(1, 1) \rightarrow (0, 1)$

2. Stability and Oscillations

   Definition 1. A network $\mathscr{N}$ is stable under a constant input, $\vec{X}_m$, if, under that input, the network, after changing, or not to a new state, will remain in said state regardless of its initial state. Otherwise, the network is said to be unstable under $\vec{X}_m$.

   Definition 2. A network is completely unstable under $\vec{X}_m$ if it is unstable for any given initial state.

   Definition 3. If a network is unstable under $\vec{X}_m$, it oscillates in one or more modes depending upon the state of the network when $\vec{X}_m$ was applied. The order of a mode of oscillation[2] is the number of states which are involved in the oscillation.

   The conditions of stability for $\mathscr{N}$-networks may be derived from their transition state matrices, $\mathscr{M}(\vec{X}_m)$, by using the following algorithm (which is a consequence of the meaning of $\mathscr{M}(\vec{X}_m)$; the proof of it is rather self-evident and has been left as an exercise for the reader):

   a.) If all terms in the diagonal of $\mathscr{M}(\vec{X}_m)$ are zero, the network is completely unstable, where the converse also holds true. Therefore, the necessary and sufficient condition for complete unstability is that the equation

$$\sum_i \mathcal{M}(\vec{X})_{ii} = 0 \tag{7}$$

has solutions. Under these solutions the network will be completely unstable.

Example 2. For the network of Example 1, the inputs which provoke complete instability are the solutions of

$$\bar{x}_1\bar{x}_2 + x_1x_2 + \bar{x}_1x_2 + 0 = \bar{x}_1 + x_2 = 0$$

which has the unique solution $x_1 = 1$, $x_2 = 0$, i. e., under the input $(1,0)$ the network is completely unstable.

b.) If any terms in the diagonal of $\mathcal{M}(\vec{X}_m)$ are 1, we delete the rows and columns which correspond to them, ending in a new matrix in which two alternatives are possible;

b_1. Some rows have only 0's.

b_2. All rows have 1's.

Under the latter case, the network is unstable when the initial state is any of those states which are present in the reduced matrix. In the former case we delete rows and columns corresponding to the states whose rows are all zeros. A new matrix is obtained that follows either alternatives b_1 or b_2. If it follows b_1, we continue the process of reduction until we end in a minor that follows b_2. If, by iteratively applying b_1, we end in only one state, the network is stable.

Example 3. The matrix $\mathcal{M}(0,0)$ for the network of Fig. XV-1 is

$$\mathcal{M}(0,0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

By deleting the first row and column, which have 1 in the diagonal, we obtain

$$\mathcal{M}'(0,0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

which follows b_2. Therefore, under the input $(0,0)$, the network is unstable if the initial state is either $(0,1)$, $(1,0)$, or $(1,1)$.

For the same network, $\mathcal{M}(1,1)$ is

$$\mathcal{M}(1,1) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

By deleting the second row and column, we conclude that the network is unstable under the input $(1,1)$ if the initial state is either $(0,0)$, $(1,0)$, or $(1,1)$. Similarly, the network

is unstable for the input $(0,1)$ if the initial state is either $(0,0)$, $(0,1)$, or $(1,1)$.

Example 4. The network of Fig. XV-2 has no external inputs. The function of each neuron is, respectively,

$$f_1 = y_1 \bar{y}_3 + \bar{y}_1 y_3$$

$$f_2 = \bar{y}_1 \bar{y}_2$$

$$f_3 = y_2 \bar{y}_3 + \bar{y}_2 y_3$$

The state transition matrix is (blanks are zeros)

		1	2	3	4	5	6	7	8
1	000	1							
2	001								1
3	010		1						
4	011					1			
5	100					1			
6	101		1						
7	110						1		
8	111	1							

The table is headed by (f_1, f_2, f_3) across the top and (y_1, y_2, y_3) down the left side.

By inspection of the diagonal, we can delete rows and columns 1 and 5. Applying criterion b_1, we then delete 8 and 4. Reapplying criterion b_1, we delete 2; again, we delete 3 and 6, ending with a single state, the 7th. Therefore, the network is stable for any given initial state.

3. Stability for a Single Neuron

For the case of a single neuron computing any of the possible Boolian functions of its M inputs, the necessary and sufficient conditions for stability adopt a much simpler form. In this case, we have only one Boolian function that describes the neuron, which is of the form:

$$y(t) = f[\vec{X}(t-1); y(t-1)] \tag{8}$$

The transition state matrix, $\mathcal{M}(\vec{X})$, is

$$\mathcal{M}(\vec{X}) = \begin{pmatrix} \overline{f(\vec{X}; 0)} & f(\vec{X}, 0) \\ \overline{f(\vec{X}; 1)} & f(\vec{X}; 1) \end{pmatrix} \tag{9}$$

Since there exist only two states, unstability of any kind implies complete unstability. Therefore, the necessary and sufficient conditions for stability are manifest in stipulating that the equation

$$\overline{f(\vec{X}; 0)} + f(\vec{X}; 1) = 0 \tag{10}$$

has no solution, i. e., the solutions of Eq. 10 produce unstability. If unstable, the neuron will be characterized by the simplest oscillation 010101... .

By negating Eq. 10, we obtain

$$f(\vec{X}; 0) \cdot \overline{f(\vec{X}; 1)} = 1 \tag{11}$$

Therefore, the solutions of

$$f(\vec{X}; 0) \cdot \overline{f(\vec{X}; 1)} = 0 \tag{12}$$

are inputs under which the neuron is stable, i. e., Eq. 12 gives the necessary and sufficient condition for stability.

Example 5. Consider the neuron of Fig. XV-3, which computes the function

$$f = x_1 x_2 \overline{x}_3 \overline{y} + (x_1 + x_2)y.$$

Equation 12 now takes the form

$$x_1 x_2 \overline{x}_3 \cdot \overline{(x_1 + x_2)} = x_1 x_2 \overline{x}_3 \overline{x}_1 \overline{x}_2 \equiv 0.$$

Therefore, the neuron is stable for any input.

R. Moreno-Diaz

References

1. W. S. McCulloch, "Agathe Tyche of Nervous Nets — The Lucky Reckoners," in Embodiments of Mind (The M. I. T. Press, Cambridge, Massachusetts, 1965).

2. R. Moreno-Diaz, "Realizability of a Neural Net Capable of All Possible Modes of Oscillation," Quarterly Progress Report No. 82, Research Laboratory of Electronics, M. I. T., July 15, 1966, pp. 280-285.

B. MYELINATED FIBERS IN THE DORSAL ROOTS OF CATS

With Dr. J. Y. Lettvin we have been studying the electrical properties of those structures in cat intradural dorsal root which we could impale with KCl-filled micropipettes. The experimental technique and preliminary results were reported in Quarterly Progress Report No. 78 (pages 281-282). We report here the conclusions reached after using two different impalement techniques.

In the first technique (tapped penetrations), the micropipette tip was positioned in a region of high resistivity by using a micrometer and the end of the micrometer barrel tapped. This normally resulted either in returning the tip to a region of low resistivity and no resting potential or in penetrating a unit showing approximately -70 mv resting potential, electrotonic charging properties, and, when sufficiently stimulated, a full-scale (>80 mv) action potential. Since the shape of the electrotonic charging curve was consistent with that expected for a parallel RC circuit, and since this would not be the case if we were charging some unit through a length of internode, we believe that the RC combination is at the point of penetration into the fiber. Similarly, we believe the point of penetration to be a point of excitability because, if it were not, the RC combination would give rise to a shunting of the action potentials from neighboring excitable regions, and this was not noticed. Hence, at the point of penetration there is an excitable membrane. That this membrane is not associated with a naturally occurring node of Ranvier can be inferred from the fact that the observed resistances (R~10 MΩ) are much too low and the observed capacitances (C~6 pf) much too high, which the frequency of penetration is much higher than would be expected if only Ranvier nodes could be penetrated. We conclude, therefore, that penetration must be internodal and that the act of penetration creates an artificial node of Ranvier whose area is somewhat larger than that of naturally occurring ones. We account for this by supposing that tapping of the micrometer barrel causes a complex orbital motion of the micropipette tip which eventually ruptures the myelin, leaving the way open for penetration of the axolemma which must be excitable along its entire length.

W. F. Pickard

C. FURTHER STUDIES WITH CYLINDER LENSES

In Quarterly Progress Report No. 73 (pages 309-316), we have described a method for testing camera lenses which makes use of the properties of the crossed-cylinder lens. In a later report (Quarterly Progress Report No. 77, pages 383-389), we showed how a small section of such a lens can be fabricated by the torsional deformation of a prism of glass. It now appears that this method can be generalized to permit the fabrication of a large-aperture sphero-cylinder lens.

The method depends on the deformation of square plates by couples of force at opposite corners. The treatment of this problem in elasticity theory may be found in standard texts; a formula that is valid for deformations that are small compared with the thickness of the plates is given by[1,2]

$$z = k \, x \, y. \tag{1}$$

We change coordinates as follows:

$$a = \frac{1}{\sqrt{2}}(x+y)$$

$$\beta = \frac{1}{\sqrt{2}}(x-y),$$

and note that

$$x^2 + y^2 = a^2 + \beta^2 = r^2.$$

Equation 1 may be transformed as follows:

$$z = k \, x \, y = \frac{k}{2}[x^2+y^2-(x-y)^2]$$

or

$$z = \frac{k}{2} r^2 - k\beta^2.$$

The first and second terms of Eq. 2 represent, respectively, a spherical component of curvature $k/2$, and a cylinder component with axes a, β and curvature $-k$. This particular combination of spherical and cylinder components is known as the "crossed-cylinder lens," and is used extensively in ophthalmic refractive diagnosis.[3] If one wishes to isolate the "pure" cylinder component, corresponding to the second term of Eq. 2, an additional spherical lens may be employed to cancel the spherical component.

The feature of this cylinder lens construction that recommends it in preference to more obvious methods, as for example, wrapping a sheet of glass around a precisely formed cylindrical mandrel of large radius of curvature, is that the forces are self-equalized and need be applied at only four points.

1. Practical Features in the Construction of Such a Lens

It will be apparent that there is a maximum refractive power, relative to the size of the plates, which can in practice be attained because of the combination of two limitations: the deformation of the plate must be small compared with the thickness of the plate, in order that Eq. 1 will apply, and the breaking strength of the glass must not be exceeded. (Quartz would give greater strength; various plastics would permit greater deformation before breaking but would have inferior optical properties.)

Let us now consider the various methods by which cylinder lenses can be fabricated from deformed plates of glass. In Fig. XV-4 we show two plates, each 3 mm thick and 78 mm square, which have been spaced with rubber shims, those in one pair of opposite corners being twice the thickness of the others, and taped to provide a reservoir for fluid or plastic. (With careful manipulation, the tape can be applied after the glass has been deformed by tightening the clamps.) The space between the plates may be filled with an inert transparent fluid of low vapor pressure such as a silicone oil, or it can be

cast with an epoxy resin for a more permanent and useful construction. For this prupose, we can recommend Maraset Type 657 A and B casting resin combination (from The
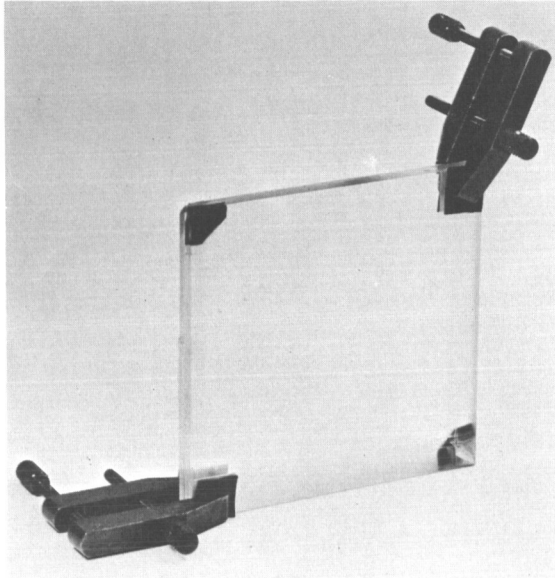

Fig. XV-4.

Two square optical glass plates, taped to contain liquid or epoxy resin, are stressed in torsion by external clamps and internal rubber shims.

Marblette Corporation, Long Island City, N. Y.) mixed as directed, centrifuged to remove bubbles, and allowed to harden undisturbed at room temperature for one week. Two out of six lenses fabricated in this way have been judged to be near-perfect; the method is still subject to some variability.

Useful results have been obtained with these plastic-cast lenses, but there are two evident disadvantages to this construction: the glass is left under strain and is thus prone to fatigue failure or breakage with small shocks; and the homogeneity of the epoxy plastic is inferior to that of optical glass, and is liable to change with absorption of solvents, age, heat, and so forth. A way around these difficulties is suggested by a study of the method used by Bernhard Schmidt to fabricate the corrector plate for his famous telescope design.[4] This method of construction, which we have not yet tested, is the following: One starts with the previous lens, epoxy resin cast between the deformed glass plates. This sandwich of glass and plastic is ground and polished flat on the outer surfaces, and the components are then separated by using heat, solvents or both. These glass components are each crossed-cylinder lenses of dioptric strength determined by the index of the glass rather than the plastic. They may be used separately, or may be glued together at their now flat inner surfaces by using the thinnest possible layer of cement, to form a double-strength, unstrained, cross-cylinder lens of superior quality.

2. Tests of Plastic-Cast Crossed-Cylinder Lens

The method that we use for testing crossed-cylinder lenses is the following: A collimator is first adjusted to project an image of a monochromatic point source. This simulated monochromatic star image is photographed with a long focal length camera lens of aperture similar to that of the collimator. Interposed between these two lens
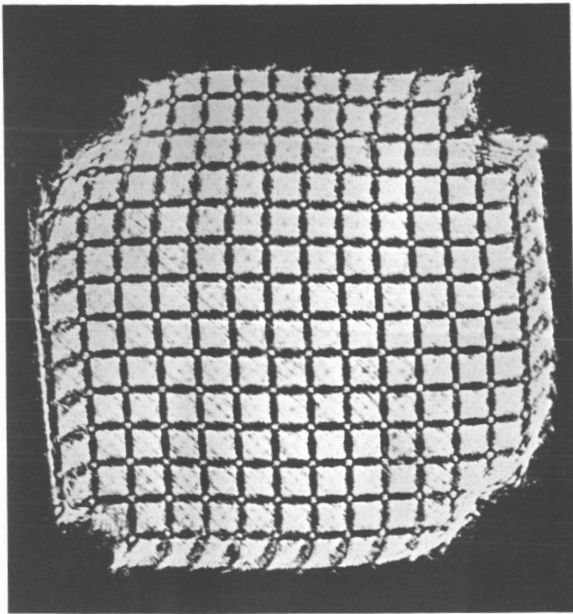

Fig. XV-5.

Test of crossed-cylinder lens of new construction; space between plates has been cast in epoxy resin. A perfect lens would image the grid without distortion.

systems is the crossed-cylinder lens, together with a rectangular grid, here a photo-chemically etched screen of beryllium copper having a 'wire' thickness that is small compared with the mesh spacing, which is 2-5 mm. The grid is oriented so as to give minimal diffraction blur, which means that its axes are parallel to the edges of the deformed square plate lens.

With this arrangement, the camera photographs what is in effect an image in miniature of the aperture of the cylinder lens as metered by the grid. The quality of the cylinder lens is indicated by the degree of distortion of the image of the rectangular grid. In Fig. XV-5 we show the results of a test of a plastic-cast lens of 78-mm square aperture, and strength ±1/8 diopter. [The optician's formula for this lens is (−1/8, +1/4 Diopters).] For this test, the angular subtense of the collimator was 1/50,000 radian, and the spectrum of the source was restricted to 5750 ± 50 Å; the camera and collimator lenses were high-quality air-spaced doublet lenses. Thus, we can reasonably ascribe the defects in the pattern of this grid, which is seen to exhibit a 'folding over' at the corners where the external clamps were applied, to the cylinder-lens construction. The
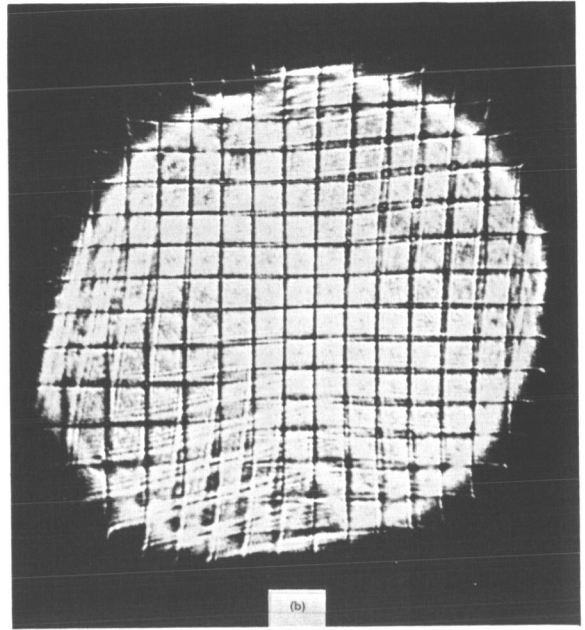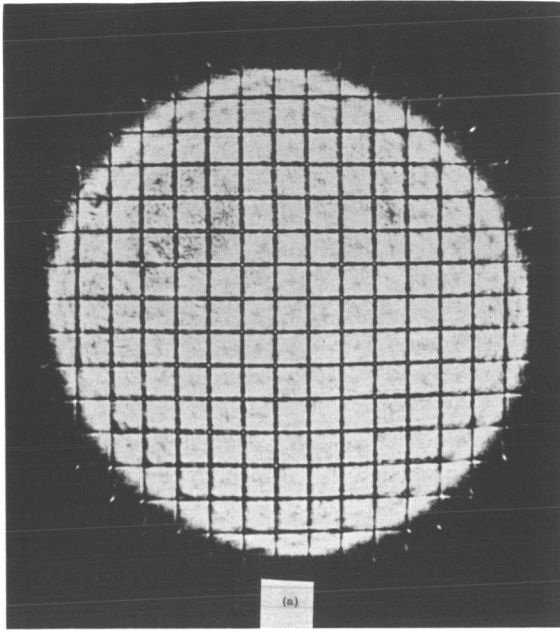
Fig. XV-6. Comparison of crossed-cylinder lenses of same dioptric powers and differing constructions. (a) Lens of new construction. (b) Ophthalmic lens.
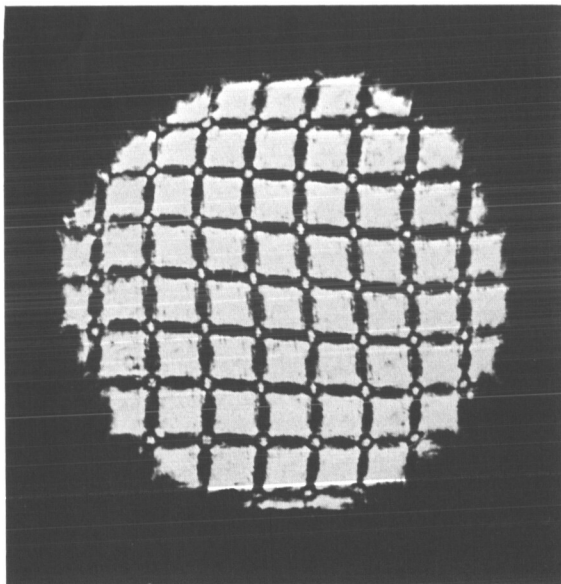

Fig. XV-7. Distortion of the pattern of the grid caused by spherical aberration.
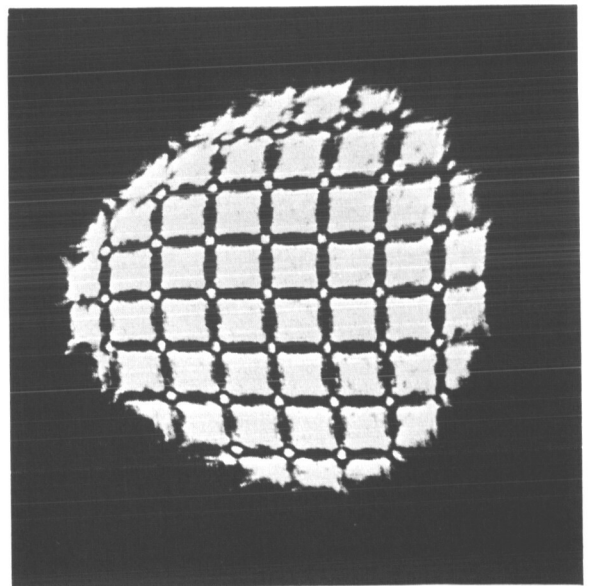

Fig. XV-8. Distortion of the pattern of the grid caused by coma.

central regions of this lens (not our best lens) are, however, unaffected by this defect.

Lenses of this type were fabricated from very good quality plate glass of 2.5-3 mm thickness; this power is the highest possible without probable breakage caused by delayed fatigue failure of the glass. By coincidence, this is also the smallest available prescription power available in ophthalmic crossed-cylinder lenses, and permits a comparison of optical quality. In Fig. XV-6a we show a test on a finer grid than before, of the central 50 mm of our very best plastic-cast cylinder lens. In Fig. XV-6b is shown a similar test of the best of three ophthalmic crossed-cylinder lenses. The improved performance of the new construction is thus demonstrated, although the comparison is obviously an unfair one. The defects of the ophthalmic lens shown here are so small that they could not be observed by any test with an unaided human eye.

3. Use of Crossed-Cylinder Lens to Test Camera Lenses

The method that we use to test crossed-cylinder lenses can also serve as a test of the camera lens, once the quality of the crossed-cylinder lens has been established. The distortion of the pattern of the grid will give qualitative indication of the predominate aberration of the lens. In Fig. XV-7 we show the pattern of the grid characteristic of primary spherical aberration, simulated here by means of a supplementary lens pair, a flat convex and a bent concave lens of opposite dioptric powers. In Fig. XV-8 we show the distortion characteristic of the off-axis aberration coma.

We regard this extension of our previously reported lens-testing procedures as being of importance, since it enables one to test the performance of the lens over its entire aperture. Thus one can detect defects of figuring, as well as formula. Still another possibility is to test optical elements, for example, mirrors, prisms, filters, window glass elements, thereby using the instrument in the manner of a striascope. The advantage is that one need not record or interpret light-intensity variations; the distortion of the pattern of the grid gives the information.

An advantage of methods that use the crossed-cylinder lens over other lens tests that make use of aperture screens (of which the Hartmann test is perhaps the best example) is that the axis of the grid can be adjusted so as to give minimal diffraction blurring of the pattern. Thus, only to the extent that the lens produces aberrations affecting the orientation of the elements of the rectangular grid, is the pattern subject to diffraction blurring. In effect, then, the more perfect the lens, or other element under test, the more sensitive is the test. In our present experiments the brightness and spot size of our collimator limit the magnification that can be attained with time exposures of reasonable duration. A neon laser of modest power should remove this limitation and enable us to determine the ultimate sensitivity of the method.

B. Howland, S. J. Wiesner

[Stephen J. Wiesner is with the Department of Physics, Brandeis University.]

References

1. A. E. H. Love, A Treatise on Mathematical Elasticity (Dover Publications, New York, n. d.), p. 471.

2. S. Timoshenko and S. Woinowsky-Krieger, Theory of Plates and Shells (McGraw-Hill Book Company, New York, 1959), Chap. 2.

3. I. M. Borish, Clinical Refraction (The Professional Press Inc. , Chiago, Illinois, 1954).

4. Amateur Telescope Making, Book Three, Albert G. Ingalls (ed.)(Scientific American, Publisher, New York, 1961), p. 365.

XVI. COMPUTATION RESEARCH[*]

Research Staff

| Martha M. Pennell | Gail M. Fratar | Elaine Isaacs |
| Heather Davis | Joan Harwitt | Eleanor River |

A. LINEARIZING THE ROOTS OF A POLYNOMIAL

Probably the most common application of the method of least squares is the determination of constants in an empirical formula whose form is to be inferred from the results of experimental data. The method, however, can equally well be used to approximate known functions by less complicated formulae. The following problem submitted to us by the Microwave Spectroscopy Group is an example of the latter. Find three sets of constants $\phi_1, ..., \phi_5$ such that one of the following expressions:

1) $\phi_1 \omega_0 + \phi_2 \omega_4$

2) $\phi_1 \omega_0 + \phi_2 \omega_4 + \phi_3 \omega_6$

3) $\phi_1 \omega_0 + \phi_2 \omega_4 + \phi_3 \omega_6 + \phi_4 \omega_8$

4) $\phi_1 \omega_0 + \phi_2 \omega_4 + \phi_3 \omega_6 + \phi_4 \omega_8 \ddagger \phi_5 \omega_{10}$

would best approximate the three positive real roots of the following sixth-degree polynomial in V in the range $0 \leqslant \theta \leqslant 45°$, $0 \leqslant \gamma \leqslant 45°$:

$$V^6 - 28.38V^4 - (47.38\omega_4 - 226.14)V^2 - (550.33 - 259.09\omega_4 + 183.55\omega_6) = 0 \tag{1}$$

$$\omega_0 = 1$$

$$\omega_4 = x^4 + y^4 + z^4 - 3/5$$

$$\omega_6 = x^6 + y^6 + z^6 - 15/11\,\omega_4 - 3/7$$

$$\omega_8 = x^8 + y^8 + z^8 - 28/15\,\omega_6 - 210/143\,\omega_4 - 1/3$$

$$\omega_{10} = x^8 + y^{10} + z^{10} - 45/19\,\omega_8 - 42/17\,\omega_6 - 210/143\,\omega_4 - 3/11$$

$$x = \sin\theta\cos\gamma$$

$$y = \sin\theta\sin\gamma$$

$$z = \cos\theta$$

Equation 1 was solved for all possible combinations of $\theta, \phi = 0, 9, 18, ..., 45°$.

Table XVI-A.

Equation Number	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
8 points					
1)	4.11098	.496842			
2)	4.11108	.496822	−.0717939		
3)	4.11130	.506908	−.0509087	−.405050	
4)	4.11137	.506378	−.0517689	−.403467	.217128
12 points					
1)	4.11129	.4960810			
2)	4.11111	.498560	−.0365156		
3)	4.11126	.507561	−.0459987	−.420735	
4)	4.11127	.507441	−.0464354	−.421370	.0686554

Table XVI-B.

Equation Number	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5
8 points					
1)	4.11037	.498241			
2)	4.11105	.496856	−.0283327		
3)	4.11123	.506867	−.0331734	−.432111	
4)	4.11108	.508400	−.0518164	−.442301	.348897
12 points					
1)	4.11108	.500282			
2)	4.11120	.500578	−.0331930		
3)	4.11126	.507069	−.0390402	−.432419	
4)	4.11128	.507350	−.0437185	−.438920	.185722

n arbitrary roots on the same roots locus were then used as the data to a least squares analysis whose fitting functions were the ω's given above. Both n and the roots used as data were varied. The results are summarized below. Table XVI-A approximates the root with the largest magnitude using 8 and 12 points. Table XVI-B repeats the same calculations of a different set of 8 and 12 points.

<div style="text-align: right">Martha M. Pennell, Heather Davis</div>

Author Index